Department of Electrical and Information Engineering
ELECTRICAL AND INFORMATION ENGINEERING
Ph.D. Program
SSD: ING-INF/05

Final Dissertation

A semantic-based biosignal mining for Affective Computing

by
Annarita Cinquepalmi :

Supervisors:

Prof. Michele Ruta

Coordinator of Ph.D Program:
Prof. Vittorio Passaro

XXIX cycle, 2014-2016

Department of Electrical and Information Engineering
ELECTRICAL AND INFORMATION ENGINEERING
Ph.D. Program
SSD: ING-INF/05

Final Dissertation

A semantic-based biosignal mining for Affective Computing

---

by

Annarita Cinquepalmi :

____

Firma leggibile e per esteso

Referees:

Prof. Umberto Straccia

Prof. Marco de Gemmis

Supervisors:

Prof. Michele Ruta

________________ *firma*

Coordinator of Ph.D Program:

Prof. Vittorio Passaro

________________ *firma*

XXIX cycle, 2014-2016

Non chi comincia ma quel che persevera

Contents

List of Figures

List of Tables

Abstract

Affective computing is the research field in charge of the computational processing of emotions. Applications range from enhancing human-computer interaction to biofeedback generation for avoiding risky situations and improving subjects' well-being.

Current automatic emotion recognizers basically provide only simplistic classification approaches and yield trivial labels using machine learning techniques without capturing relations between biosignals and observations measured by the various sensors.

This thesis proposes a framework for affect recognition monitoring human body vital signals through wearable, non-invasive sensors. The intrinsic complexity in emotion detection is tackled by means of automatic discovery based on a semantic matchmaking process via non-standard reasoning services. In particular, manipulation of vague concepts such as emotions and their dynamic evolution are achieved exploiting Fuzzy Description Logics (DLs) ontology-based approach.

The proposed architecture is able to receive in input time series of biosignals, extract meaningful high-level knowledge from mining, identify emotional patterns through the flexibility of Fuzzy-DLs, and exploit semantic-based matchmaking to recognize user emotions. Prototypes were implemented w.r.t. a reference dataset and preliminary experimental tests were carried out to verify the feasibility of the approach on emotions experienced by users. The system enhances human-computer interaction allowing a feedback generation for subjects and improvement of their well-being.

Motivations for the work stem from the idea that a revision of existing approaches can have a significant impact on the effectiveness and applicability of affective computing, by combining logic-based Knowledge Representation (KR) with machine learning techniques in a low-cost wearable computing set-up. The social aspects of improved Affective Computing systems concern mainly the potential applications in diagnosis, treatment and management of mental and stress-related disorders. The possibility to support biofeedback in users to oppose undesirable emotions and behaviors can also have a significant impact on substance abuse and other unhealthy habits, affecting the quality of life of the general population and welfare policies.

Chapter 1

Introduction

According to R. Picard [98], *Affective Computing* (AC) aspires to create computational systems which are "emotionally intelligent" (EmI), *i.e.*, capable to recognize, understand and express emotions in order to improve users' well-being. EmI systems may establish empathy with the user *e.g.*, through an *affective avatar*, *i.e.*, an interactive automated agent designed to perceive user emotional experiences when engaged in specific activities.
Physiological signals have been used increasingly in AC thanks to technological improvements in low-cost miniaturized unobtrusive wearable biosensors for continuous monitoring. Recently, manufacturers have being developing increasingly robust and cost-effective biosensors for fast and sensitive analysis of human body vital signals. Over the last decade, EmI systems have gained momentum for a wide number of applications in several important companies: NeuroFocus[1] utilized electroencephalography (EEG), eye tracking, and biometrics to capture the non-conscious aspects, emotions, and preferences of consumer decision-making; EmSense[2] developed the proprietary unobtrusive EmBand$^{\text{TM}}$hardware for measuring positive/negative emotional response and cognitive engagement to advertising. Emotion-aware systems identify specific outcomes from biosensors and respond by triggering appropriate actions within a given context. Additional examples include: (i) monitoring the elderly to recognize signs of health issues [120], such as sadness bouts as a symptom in depressed patients, and alerting healthcare providers; (ii) increasing safety of drivers by observing their emotions [50] and, suggesting a relaxation technique if a state of anger or frustration is detected (*biofeedback*); (iii) improving user satisfaction in smart home environments [26], by controlling domotic devices to favor comfort and resting. Literature evidences the usefulness of biosignals analysis in emotion detection and classification.

[1]www.neurofocus.com
[2]www.emsense.com

Nevertheless, most existing approaches are still quite intrusive. Furthermore, studies are tipically carried out in controlled laboratory conditions, hardly transferable to real scenarios. Basically they rely on conventional computing architectures running procedures for signal processing and features extraction which have high computational costs, affecting the performance in real-time applications.
However, several practical questions arise when dealing with Affective Computing system: 1)What does it mean for an emotional system to express emotions that it doesn't feel? 2)Once the AC system has sensed the user's biosignals, how it understand emotional patterns? 3)How the system differ from current emotional classification software? and 4)How the system respond to user's emotional state? To answer all these questions, this thesis proposes a quasi-real-time computing framework which only leverages off-the-shelf technology for biosignal monitoring and analysis, attempting to go beyond trivial emotion classification. Intrinsic complexity in emotion detection is coped with by exploiting Knowledge Representation, and particularly standard Semantic Web languages enriched according to the Fuzzy Description Logics (DLs) formalism [122]. Biosignals and features are described through semantic annotations based on a reference ontology. In particular, Fuzzy DLs enable the description and manipulation of vague concepts such as emotions. Semantic-based mining of raw sensor data makes them machine-understandable and allows for knowledge to be processed efficiently, even in mobile and pervasive contexts with severe resource limitations in memory, storage and energy consumption. For this purpose, the optimized Mini-ME[3] embedded reasoning engine [116] is adopted.
The framework works in three fundamental stages: (i) detect most relevant biosignal features; (ii) build an annotated description of the emotional dimensional model in terms of *valence* (V) and *arousal* (A) [109] via Fuzzy DL-Learner[4] [11]; (iii) use matchmaking to recognize the emotion from its dimensional features. Non-standard inference services [112] are exploited to compare annotations of the VA space, discovering the most emotion(s) experienced by the user. The ultimate goal of the framework is to provide helpful and timely user feedback and/or provide customized services. The free public DECAF data set [1] was used for the initial implementation and experimentation of the proposed framework, in order to build a Fuzzy DL Knowledge Base (KB) of emotions correlated with biosignals and the continuous VA model. Experiments on the proposed framework provide an early validation

[3]Mini-ME - the Mini Matchmaking Engine. http://sisinflab.poliba.it/swottools/minime

[4]Fuzzy DL-Learner. http://www.umbertostraccia.it/cs/software/FuzzyDL-Learner/index.html

on the feasibility of emotion identification by combining logic-based Knowledge Representation (KR) with fuzzy machine learning techniques.
The remainder of the thesis is organized as follows. Chapter 2 provides a literature analysis on Affective Computing discussing both biosignals and knowledge representation, while Chapter 3 describes the proposed framework in detail. In order to verify the validity of the proposal, experiments are reported in Chapter 4. Finally, in Chapter 5 the thesis closes with concluding remarks and future perspectives.

Chapter 2

Background

2.1 Biosignal-based affective computing

Emotion plays an important role in human experience, influencing cognition, perception, learning, communication, and even rational decision-making [94, 92]. However, there is no universal definition of the term emotion. According to [84] human emotion involves *"...physiological arousal, expressive behaviors, and conscious experience"*. Emotional and social users' needs are involved in Human-Computer Interaction (HCI). The general communication theory named *Media Equation* asserts that people treat computers and other media as if they were either real people or real places [106]. In response to the new challenge of HCI, a new field of Artifical Intelligence -*Affective Computing*- has emerged. The term was coined by its pioneer Rosalind Picard at the Massachusetts Institute of Technology in the following way:

"Affective Computing is the research area concerned with computing that relates to, arises from, or deliberately influences emotion." [94]

The functionality a computer should have to fulfill Picard's description have been synthesized by Hudlicka [54] as shown in Figure 2.1:

1. Affect Sensing and Recognition. The first step towards an automatic emotionally intelligent machine is to observe the user's affective states and the transition beetween them.

2. Adapting to user affect. An affective computer after assessing the human's emotional state should translate it into patterns, in order to create the computer's affective model and respond appropriately to the user.

3. Machine "affect expression". It is an open issue whether a machine can have an affective state, but machines can behave as if reflecting a particular emotion. For example, an affective avatar can support a user while performing a task.

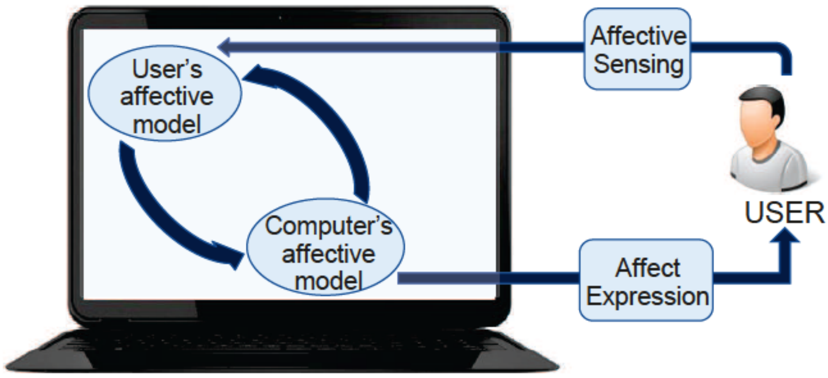


Figure 2.1: Human-computer interaction in Affective Computing.

4. Modeling affect in user and machine. The ability to decide how, or whether, to respond to the user's affective state, can also help disambiguate sensor data (*e.g.*, high skin temperature in positive emotion *vs.* high skin temperature in negative emotions), or compensate for lack of data, and help establish the best means of responding to a specific affective state. To accomplish this, the affective machine must merge current stimuli with contextual information about the user and the environment where it is – *e.g.*, user's past experiences and cognitive abilities, current goals and needs – to deduce the most probable affective state, which then directs subsequent behavior.

Over the last decade, Emotional Intelligence (EmI) systems have gained momentum for a wide number of applications: emotion-aware systems identify specific outcomes from biosensors and respond by triggering appropriate actions within a given context. The main application areas are:

- **E-learning**: help learner when he/she is performinng a task and if the learner is fidgeting or frustated, and repeats the same mistake, then the affective machine can suggest a different strategy or pass control over to human operators [57, 97, 108].
- **Automotive**: increase safety of drivers by observing their emotions and, if a driver is in a state of frustration, the affective machine suggests a relaxation technique (*biofeedback*) [50, 87].
- **Games**: perceive user emotional experiences when engaged in gaming activities and adapt the difficulty level to maintain the player engaged when the machine detects him/her getting bored [19, 77, 105].
- **Neuromarketing**: capture the non-conscious aspects, emotions, and preferences of consumer decision-making; for example, NeuroFocus used

electroencephalography (EEG), eye tracking, and biometrics, and EmSense developed the proprietary unobtrusive EmBand^{TM}hardware for measuring positive/negative emotional response and cognitive engagement to advertising.

- **Healtcare**: monitor the elderly to recognize signs of health issues [20, 120], such as sadness bouts as a symptom in depressed patients, and alerting healthcare providers. Similarly, detect subject's emotional state during a rehabilitation task can improve its effectiveness and patient's interest. Furtermore, the system can assign personalized therapies [12, 59].

- **Psychology**: detect emotions of autistic people in order to improve human communication and as a complementary part of conventional therapies [9].

Other examples include monitoring stress in ambulance dispatchers [83] or fall detection in the elderly [68]; security authentication and information sharing [21]; improving user satisfaction in smart home environments [26], *e.g.*, by controlling home environment actuators in order to favor resting of a stressed user and finally displaying participants' emotions during a conference to enhance audience social interaction [28].

2.1.1 Basics

User emotions and affective states can be collected in multimodal ways: through facial expressions, speech, body gestures, vocal intonation and physiological signals. A brief description follows:
Facial expressions. Techniques for emotion recognition basically segment facial images in regions of interest, typically mounth, eyes, eyebrows, chin and wrinkles. Authors in [24] discussed how computer vision techniques can be employed for measuring and encode facial expressions and attempt to classify them into six basic emotions, focusing on the physical transformations of the face or the interpretation of the expression corresponding to a certain internal emotional state.
Speech. Emotion recognition from vocal signals focuses on the analysis of speech paralinguistic features, such as prosody and intonation. Authors in [16] discussed how a speech-based affect detector requires acoustic and prosodic features extraction process. Unfortunately, most of the automatic recognition systems take as input vocal streams uttered by actors in laboratory controlled conditions and stored in databases rather than spontaneous expressions, which are more difficult to analyze. The analysis is based on

a combination of words that explicitly express the emotional state and the tone emitted when different emotional states are developed [129].
Body gestures. Emotional features can be found analyzing body posture and movement. In particular, body configurations identify specific affective states. Kipp and Martin [61] examined the relation between emotion and gestural features on actors in filmed theater stagings. One of the major drawbacks is that most of these studies are not able to recognize emotions independently of the action the person is doing, but relied on a limited set of acted body expressions (*e.g.*, dance, gesture, posture, and gait) [63].
Biosignal. Physiological changes continuously co-occur with emotion. They can be explored by studying the relationship between the autonomic nervous system and emotions. Affective machines that automatically sense these changes and identify patterns between emotion and physiological responses are still an open challenge for researchers. Healey *et al.* [50] determined a driver's stress state during real driving tasks by measuring his physiological signals. In the last few years, research provided wide support for the assertion that affective recognition systems using biosensors can be standardized following these guidelines [98]:

- **Signal acquisition**. An initial data acquisition process collects physiological signals from human body via sensing devices.

- **Signal preprocessing**. Gathered raw data are always contaminated with noise and artifacts due to measurement techniques and human muscular movement. Before processing, it is necessary to use filtering systems on raw data in order to make them readable and meaningful. Low-pass and smoothing filters are usually applied to raw data and the baseline signal component is subtracted from the entire time series. The baseline corresponds to the signal measurement acquired without any emotional stimulus. Normalization is often part of preprocessing as well.

- **Feature extraction**. Once the signals are preprocessed, meaningful characteristics have to be extracted. They include statistical (*e.g.*, mean, standard deviation), time-domain (*e.g.*, rise and recovery time) and frequency-domain features, or more complex ones (multi scale sample entropy, sub-band spectra) [47].

- **Emotions classification**. After selecting the relevant features, they must be used to train a classifier in order to discriminate emotional states. A vast set of *off-the-shelf* machine learning methods has been applied to develop emotion recognizers. These include Linear Discriminant Analysis, Neural Networks, K-Nearest Neighbors, Support Vector Machines, Decision Trees, Bayesian Networks and Fuzzy Rules [47, 60, 86, 98, 104].

2.1.2 Emotion modeling approaches

Emotion theories can be classified into three categories:

- **Evolutionary**. They were introduced by Charles Darwin, who applied to human communication his theory based on evolution and natural selection [30]. According to Darwin, human emotions exist because they attend an adaptive role to environmental stimuli.

- **Physiological**. They focus on physiological reaction induced by emotions [55]. In particular, the theory suggests the emotional reaction is a consequence of human physical reactions and not their cause.

- **Cognitive**. Cognitive psychology regards emotion as a mental process which stems from a cognitive unconscious process called appraisal. Researchers argue on the connection to the appraisal and the personal emotional experience. The brain is the main organ involved in processing and evaluating emotional events [71].

Several cognitive models of emotions exist. The most commonly used ones in affective computing are the discrete *categorical* model, introduced by Ekman [35], and the continuous *dimensional* one, proposed by Russell [111]. The categorical model describes emotions using separate and distinct emotional words or *affective labels*. Ekman [35] claimed some emotions are basic and biologically universal: happiness, sadness, surprise, anger, fear, disgust. These basic emotions are recognizable from human facial expressions in all cultures. A main issue in the categorical model is which emotions should be considered as label. The major partition proposed by theorists divide emotions into *primary* and *secondary* groups. Primary emotions are those affective states common to all people and possibly even animals, as asserted by Ekman [36]. A secondary emotion is generally a refined version and a specialized form of basic emotions [90]. Figure 2.2 shows a hierarchical emotional topology proposed by Parrott, emotions in the biggest circles are the six primary emotions, while all other emotions are specialised forms. A further issue is the emotions label translation in differernt languages is not the same, *e.g.*, disgust does not have an exact translation in Polish [110]. In the dimensional model, dimensions represent the essential aspects of emotions as continuous values. They have been used for real-time monitoring in

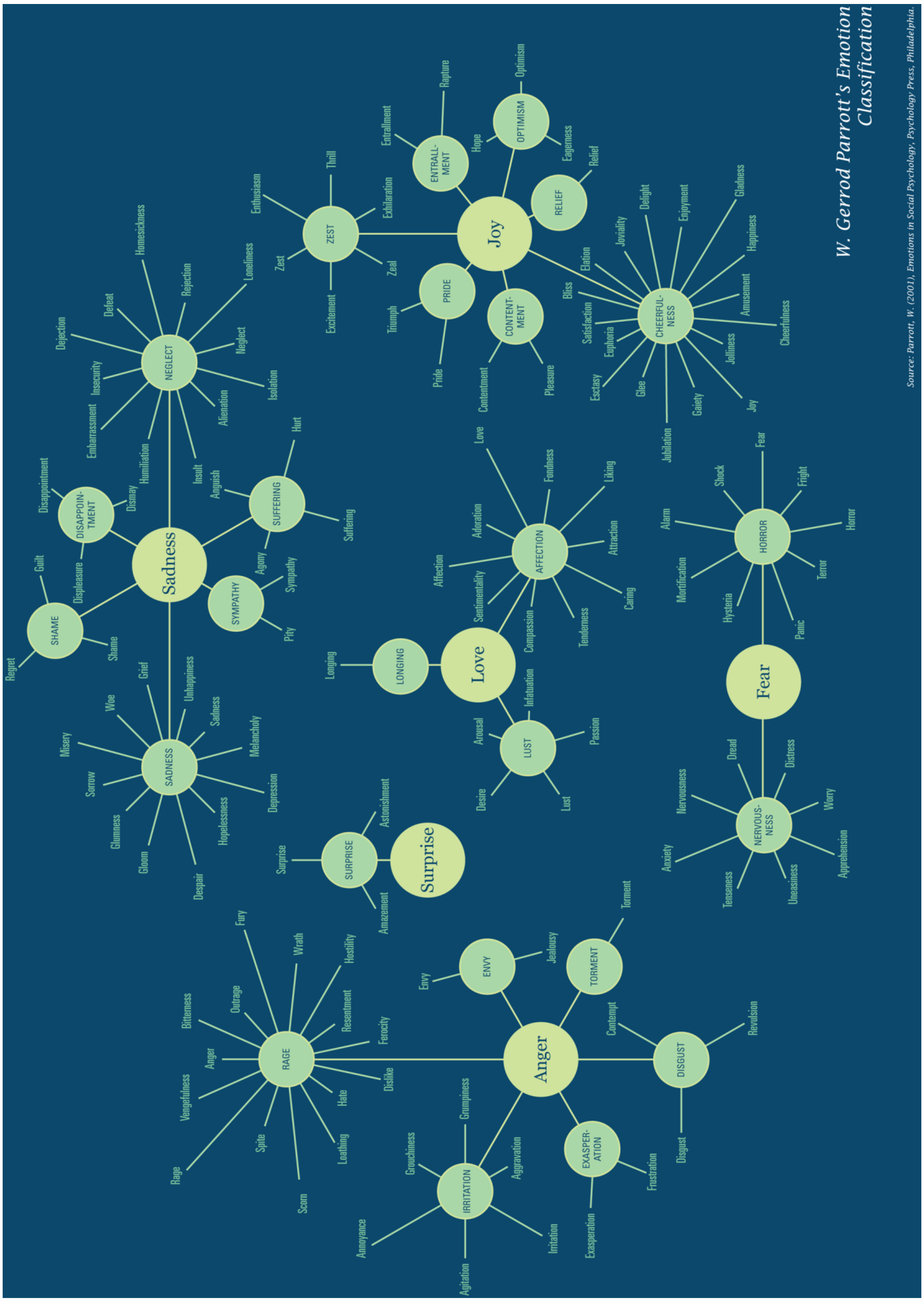


Figure 2.2: Parrott's Classification of Human Emotions [90].

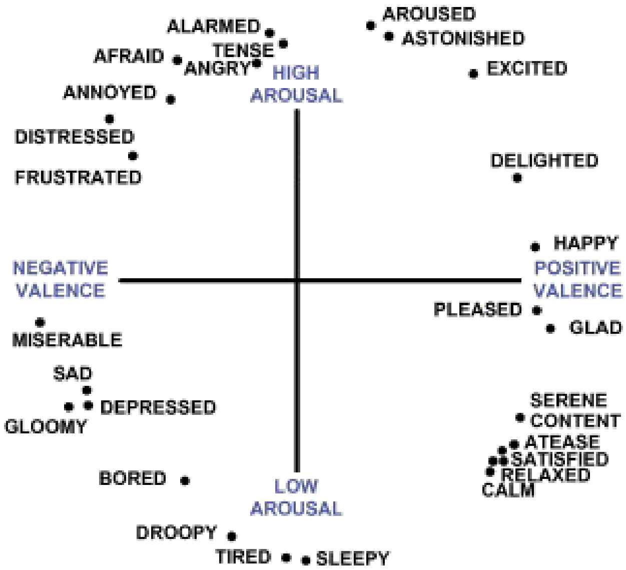


Figure 2.3: Russell emotional mode [109].

the emotion assessment process. A simple yet effective theory [109] employs a bidimensional emotion model defined by two basic parameters: *valence* or *evaluation* and *arousal* or *activation*. Valence represents how positive or negative (*i.e.*, pleasant or unpleasant) an emotion is, while arousal represents a passive/active scale, ranging from calm to excited. The key idea in the above model is to represent emotions with just a coordinate system conveying basic attributes. As a consequence, any emotion could be represented as a point in this space [109] as depicted in Figure 2.3. Every emotional state can be defined as a combination of these dimensions, *e.g.*, anger can be characterized by high arousal and negative valence, happiness by low arousal and positive valence, and sadness by low arousal and negative valence. If the emotion is completely neutral with respect to the emotional dimensions it should be assigned to the middle point of the scale. The results in [111] confirmed this space is an effective way to represent emotions interculturally, as emotions were mapped from English, Estonian, Greek and Polish subjects.

A drawback of the dimensional model is that the projection could not be unique, *i.e.*, two or more labels could have the same coordinate in the valence-arousal space. For instance, fear and anger are two highly negative and aroused emotions. To ensure a more complete emotional state description, a third dimension was added, *dominance* or *power* [82]. Dominance is related to the amount of control a subject has in a certain situation, in order to

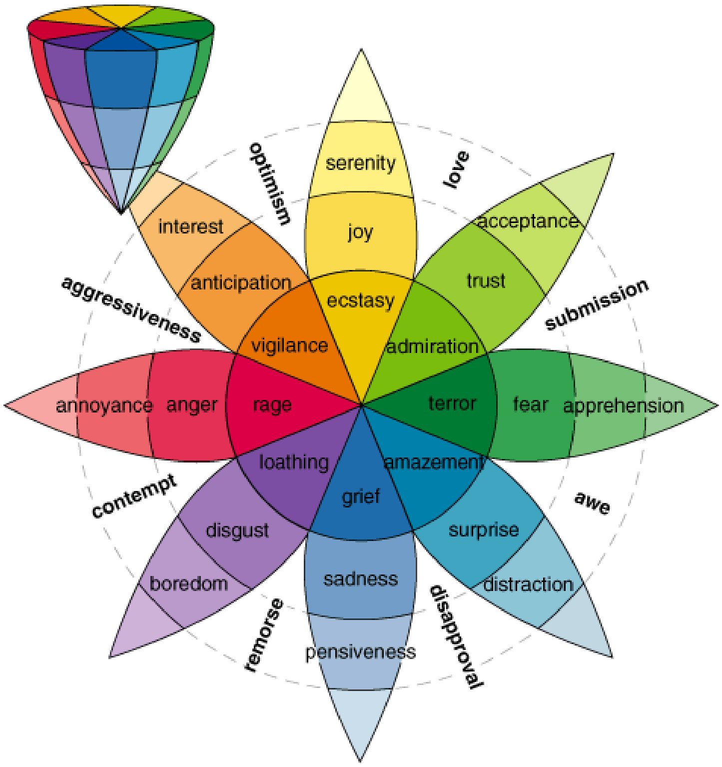


Figure 2.4: Plutchik two/three-dimensional circumplex model [100].

distinguish between emotions connected to reactions of approach and forcefulness from the ones linked to withdrawal and disengagement. In the above example, when a person is feeling fear she does not have control over the situation, while she does with anger.
Plutchik [100] combined the dimensional and categorical models in a ciecumplex, *i.e.*, wheel, using colors to show the different intensity of eight basic emotions organised into four opposing pairs (*e.g.*, ecstasy vs. grief). The eight emotions can be mixed together to form secondary emotions as shown in Figure 2.4.

2.1.3 Biosignals

Biosignals are multichannel time-varying recordings of parameters controlled by the central and/or the autonomic nervous system. They are known to convey information which can be used for emotion assessment [64]. Using biosignals has advantages over other methods:

- they are relatively robust to voluntary control and faking, because they are directly controlled by the human autonomous nervous system;
- with the proper tools, they can be gathered anytime and anywhere without active user input;
- it is possible to continuously gather information about the emotional changes of users while they are connected to biosensors;
- biosignals can be easily correlated with external channels like facial expression.

Different biosignals recognize different emotional states. A short presentation of the main types follows.

- **Electrocardiogram (ECG)** measures the electrical activity of the heart. A normal signal has a characteristic shape composed by 5 main waves [23]: a *P wave*, a *QRS complex* and a *T wave*. Relevant features extracted from ECG include heart rate (HR), inter-beat intervals (IBI) and the heart rate variability (HRV). Increase or decrease of HR can be associated with different emotions: Rainville *et al.* [102] observed an increase of mean HR for anger, fear, happiness and sadness compared to a neutral state. Ekman *et al.* [37] found HR was able to separate happy, disgusted and surprised emotional states from angry, fearful and sad ones. Concerning HRV, an energy shift from high to low frequencies is known to be associated with withdrawal reactions determined by the parasympathetic nervous system [7]. Finally, an energy decrease in the high frequency bands was observed in [102] for fear and happiness compared to neutral states, while it was significantly different for fear and anger.
- **Electroencephalogram (EEG)** measures voltage fluctuations resulting from ionic current flows within brain neurons. The brain is divided into four lobes according to the names of the bones that surround them and the sulci that separate them: the frontal, parietal, temporal and occipital lobes. Each of those lobes has been considered to be specialized in particular cognitive tasks. The brain area which is assumed to play a key role in the elicitation of emotions is the prefrontal cortex [31]. This area is located at the front of the frontal lobe and contains the orbito-frontal cortex, which is known to be involved in high-level cognitive processes such as decision making. Davidson [31] showed the left prefrontal cortex is involved in approach reactions while the right prefrontal lobe is more involved in withdrawal reactions.

- **Skin temperature**, defined as the temperature of the fingertip, is a parameter which provides information about the autonomic nervous system. Activation of the sympathetic system leads to vasoconstriction in the extremities, and as such to lower skin temperatures. Ekman and Levenson [37] showed average increases of finger temperature between 0.1°C and 0.2°C due to anger, while for fear finger temperature decreased between 0.01°C and 0.08°C. Moreover, normal subjects can voluntarily increase or decrease their finger temperature by 0.5°C and up to 1°C, when receiving temperature biofeedback [41].

- **Respiration**: girth sensors measure how deep and fast a person is breathing. Breathing pace and depth alter in reaction to an increase in heart rate and perspiration. Slow breathing rate is linked to relaxation while irregular rhythm, fast and deep breathing, quick variations, and cessation of respiration correspond to more aroused emotions like anger or fear, but sometimes also joy. Rapid shallow breathing can indicate tense anticipation including panic, fear or concentration. An overview of how the breathing rhythm is related to emotions w.r.t. brain processes was given by Homma and Masaoka in [52]. They also found an increase in the breathing rate during anticipatory anxiety which was unrelated to oxygen consumption.

- **Galvanic Skin Response (GSR)** or electrodermal activity (EDA) refers to changes in the skin ability to conduct electricity. These changes have been repeatedly proved to correlate with emotional changes [118]. GSR increases linearly with a person's level of arousal or cognitive workload, especially when the skin is sweaty [69]. The sweat glands are innervated by the sympathetic chain of the autonomic nervous system, so EDA was said to reflect sympathetic activation. The *Skin Conductance Level (SCL)* feature denotes the slowly changing part of the EDA signal. Although it can vary widely among different subjects and within the same subject in different psychological states, the typical range is between 2 μS and 20 μS [15]. It is common for SCL to gradually decrease while subjects are at rest, rapidly increase when novel stimulation is introduced, and then gradually decrease again after the stimulus is prolonged or repeated.

- **Blood Volume Pressure (BVP)** is a measure to determine the amount of blood currently running thought a vessel. When a person is startled, fearful or anxious the envelope of the BVP signal contracts; otherwise, in relaxing case the amplitude increases because there is greater blood flow to the extremities [95].

- **Electromyogram (EMG)** is the muscle electrical activity signal. The action potentials generated by the somatic nervous system travel along the muscle fibers and lead to muscle contraction. Different types of emotional imagery can elicit different patterns of EMG facial activity over the brow (*muscle corrugator supercilii*), cheek (*zygomaticus major*), and around the mouth (*depressor anguli oris*) muscle regions [115]. The EMG signal amplitude is typically between 1 μV and few mV [5]. EMG is correlated with negatively valenced emotions because high muscle tension often occurs under frustration [85]. Table 2.1 shows facial EMG activity among different basic emotions according to Ekman and Friesen [34].

- **Electrooculogram (EOG)** is the signal obtained from eyelid closures blinks. Spontaneous eyeblink identification was based on slope, amplitude and temporal criteria of the vertical EOG [121]. The blink rate can be affected by the emotional state of a person, particularly, spontaneous blink rate inhibition was considered an index of increased attention during visual tasks [39]. A reduction in blink rate is commonly observed during tasks involving visual engagement, or requiring external information processing [121].

2.1.4 Biosensors

Autonomous and unobtrusive sensor nodes or Body Sensor Unit (BSU) can form a Body Area Network or Body Sensor Network (BSN) [70]. BSUs are typically worn on the body or they are integrated into common everyday life objects (*e.g.*, shoes, shirts, earrings, computer mouse), or they are even implanted within the body. The specific application requirements govern the type, position on the body and number of sensors. Sensor devices provide unobtrusive, continuous acquisition and real-time feedback to the user. Two types of wireless nodes can be distinguished: sensors and actuators.

- *(Wireless) sensor node*. A device which wirelessly gathers and process acquired data, *e.g.*, body temperature, galvanic skin response.

- *(Wireless) actuator node*. A device that acts according to received data or through user interaction, *e.g.*, an actuator equipped with a built-in reservoir and pump administers the correct dose of insulin based on the glucose level measurements.

The main kinds of sensor nodes are *physiological sensor*, which detect the phisiological changes in human body, and the *biomechanical sensors*, which

Table 2.1: Predominant facial actions during the expression of elementary emotions [34].

Elementary emotions	**Muscles involved**	**Produced actions**
Happiness	- Orbicularis oculi - Zygomaticus major	- Closing eyelids - Pulling mouth corners upward and laterally
Surprise	- Frontalis - Levator palpebrae superioris	- Raising eyebrows - Raising upper eyelid
Fear	- Frontalis - Corrugator supercilii - Levator palpebrae superioris	- Raising eyebrows - Lowering eyebrows - Raising upper eyelid
Anger	- Corrugator supercilii - Levator palpebrae superioris - Orbicularis oculi	- Lowering eyebrows - Raising upper eyelid - Closing eyelids
Sadness	- Frontalis - Corrugator supercilii - Depressor anguli oris	- Raising eyebrows - Lowering eyebrows - Depressing lip corners
Disgust	- Levator labii superioris - Levator labii superioris alaeque nasi	- Raising upper lip - Raising upper lip and wrinkling nasal skin

measure user movements and the physical body position. In detail physiological sensors include:

- Heart Rate Sensor. Heart rate sensors detect a subject's HR in bpm (beats per minute) and HRV. These parameters can be recorded by placing an array of electrodes either directly on the surface of the chest or alternatively on the limbs. Notice that placing electrodes on the limbs is less unconfortable but more sensitive to signal artifacts [103]. In medicine, a standard electrocardiography considers 12-lead (electrode) configuration; otherwise for affective computing applications the number can be reduced to a 2-lead configuration when only the HR is being measured. Other recording methods include chest/wrist bands or finger clips, as shown in Figure 2.5.

- Electroencephalography. EEG is recorded by multiple electrodes placed on the scalp, measuring voltage fluctuations resulting from ionic current flows within brain neurons. The signal measured by an EEG electrode is thus the combination of the signals produced by all the neurons that are close enough to be recorded. The trace obtained from an electroencephalograph consists of waves of various voltage and frequency. The frequency range -in cycles per second (Hz)- of brain waves can be recognized from 0.5 to 500 Hz [125]. Important frequency ranges that are clinically relevant are divided into four main groups: *alpha*, *beta*, *theta* and *delta* bands. The frequency of EEG measurements tipically ranges from 1 to 80 Hz, with amplitudes of 10 to 100 in μV [56] characterizing EEG waves. Notice that the frequency bands associated to each rhythm can vary from a study to another and from one person to another, so in the normal EEG trace all these waves are present mixed together in different proportions depending on prevailing mental activity. Along similar lines, [27] argues brain areas are differently synchronized depending on the type of stimulus. The reason why those rhythms are observed and fluctuate is still unclear, but it is supposed that synchronization of several populations of neurons could be the source of these phenomena [93]. The electrode number varies from 3 to more than 256 and depends on the measured locations on the head and the required spatial resolution. Figure 2.6 depicts a standard 10-20[1] EEG array and a 256 lead array.

[1]The "10" and "20" refer to the fact that the actual distances between adjacent electrodes are either 10% or 20% of the total $front - back$ or $right - left$ distance of the skull

- Respiration Sensor. Respiration sensors measure breathing rate using an elastic chest band as shown in Figure 2.7. The embedded piezoelectric sensor measures the expansion and the contraction of the band in order to calculate respiration rate and volume.

- Electrodermography. The skin electrical properties vary depending on the sweat level and the associated eccrine glands. The skin resistance variation is measured by passing a small fixed voltage across the electrode. Galvanic skin electrodes are placed across the hand's palm or an electrode on each of two fingers (usually middle and index finger) as depicted in Figure 2.8. EDA measurements can be influenced due to environmental conditions and user's physical activity so it is good practice to measure also user's motion, ambient humidity and temperature to limit artifacts [64].

- Photoplethysmography. This technique measures BVP signal. It consists of a light source from infra-red LED and photosensor attached to the skin for measuring the amount of light trasmission/reflection through/from the skin blood [2]. It is typically measured with the sensor placed on the fingertip as shown in Figure 2.9.

- Electromyography. Surface electromyography measures electrical potentials occurring on in the skin when the underlying muscle contracts, *e.g.*, on the face, arm or leg. Suggested electrode placements for surface EMG recording of the facial muscles, based on [42], are depicted in Figure 2.10.

- Electrooculography. This technique measures the resting potential of the retina. It is mainly used for eye movement analysis. The electrooculogram (EOG) is measured between two electrodes attached at the right and left side of the eye (for the horizontal eye movements) or below and above the eye (for the vertical eye movements). In the vertical EOG, blinks are visible.

An overview of the main biosensors characteristics is summarized in Table 2.2.

As shown in Figure 2.11, physiological parameters gathered through multiple wireless links are passed on to a WBAN coordinator –also located on the body– which routes them to the Body Central Unit (BCU). In fully wearable solutions the BCU is a portable device (*e.g.*, a smartphone) with constrained computational resources for real-time DAQ (Data Acquisition).

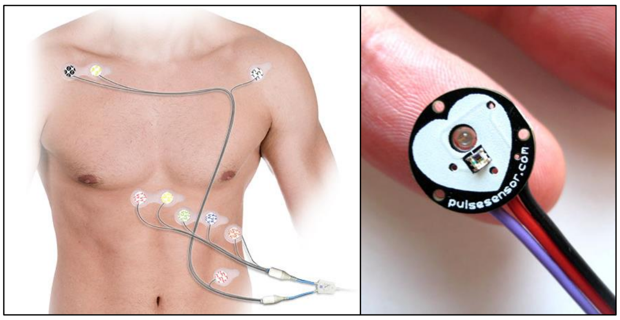


Figure 2.5: ECG sensors.

Figure 2.6: Electroencephalography.

Figure 2.7: Respiration chest band.

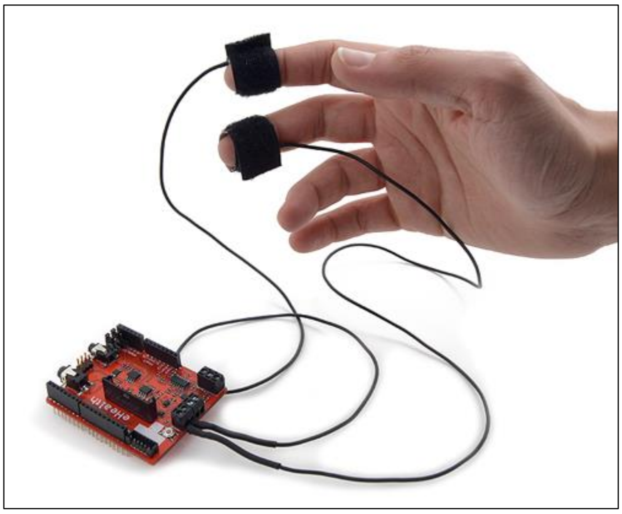


Figure 2.8: GSR sensor.

Figure 2.9: Photoplethysmography.

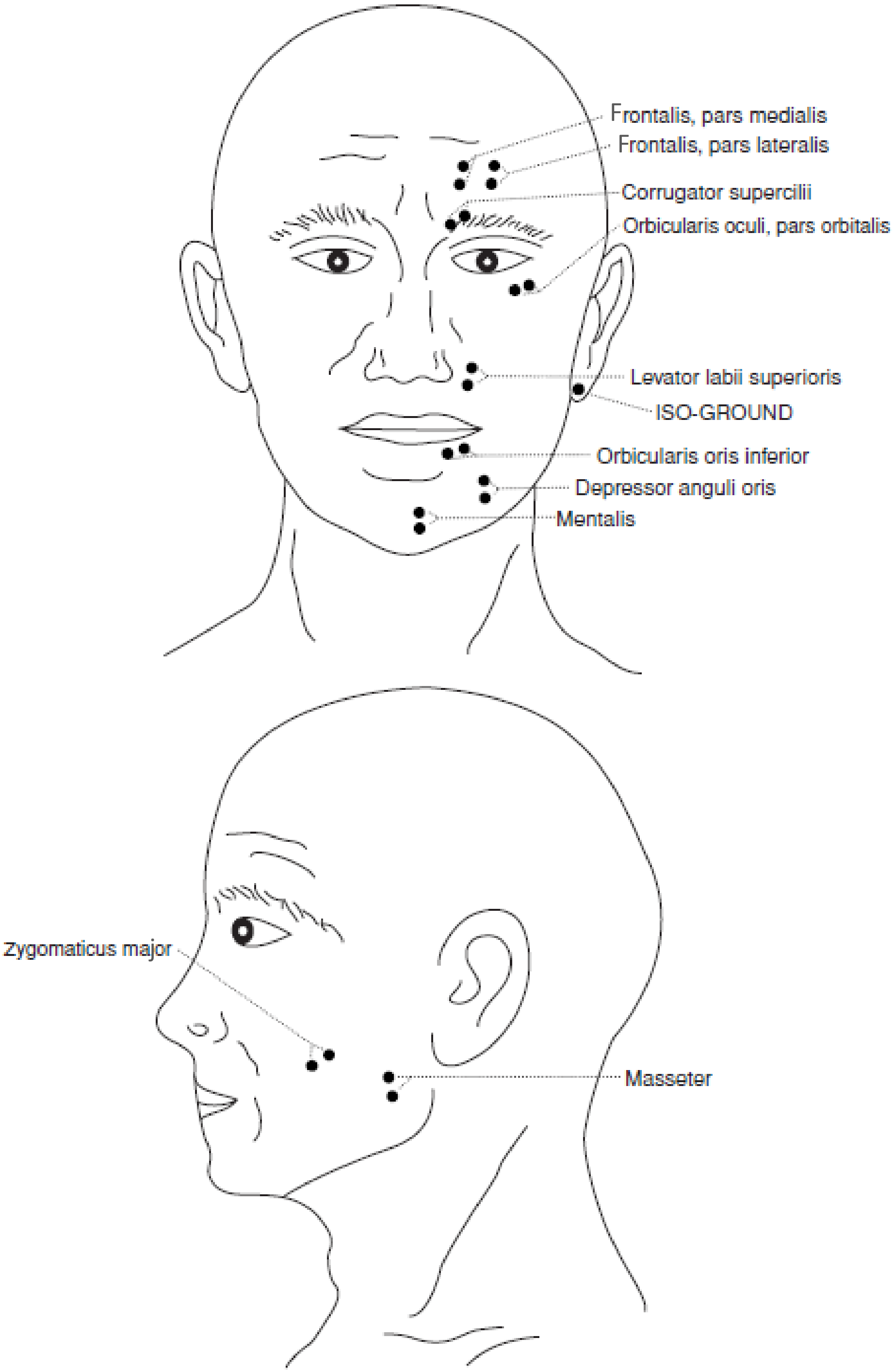


Figure 2.10: EMG electrode placements [42].

IEEE 802.15.6[2] is the international standard developed by IEEE 802.15 Group 6 for BANs for short-range, low-power and high-reliability wireless communication for use in close proximity to, or inside, a human body. In [67] authors describe advantages with respect to other wireless standards communication protocols. The main BANs characteristics are [67]:

- *short range*: BSUs communicate with each other in short range (2-5 meters);
- *network density*: people can have different body area networks on them. BAN standard allows 2-4 networks per m^2;
- *network size*: BAN standardization group expects a maximum of 256 devices per network;
- *data rate*: very different data rates depending on the type of applications (from 10 kB/s to 10 MB/s);
- *power consumption*: ultra low power consumption from 0.01 mW to 40 mW in order to protect human tissue;
- *lifetime*: in-body sensors should have a long lifetime (one week or longer);
- *low cost*: biosignals should to be monitored inexpensively;
- *low complexity*: sensors should be easily produced and implemented;
- *security*: IEEE 802.15.6 defines three security levels: level 0 - unsecured communication; level 1 - authentication only; and level 2 - both authentication and encryption.

2.1.5 Affective computing datasets

Affective datasets are costly to acquire. Table 2.3 summarizes the experimental setup, acquisition modalities, storage and biosignals of freely available datasets. Despite the large amount of studies on emotions, there are only few public available affective databases which include physiological signals and audio-visual data [33, 45, 81, 89]. They mostly include speech in different languages, face and body gestures or audio-visual data.

[2]IEEE WPAN Task Group 6, http://www.ieee802.org/15/pub/TG6.html

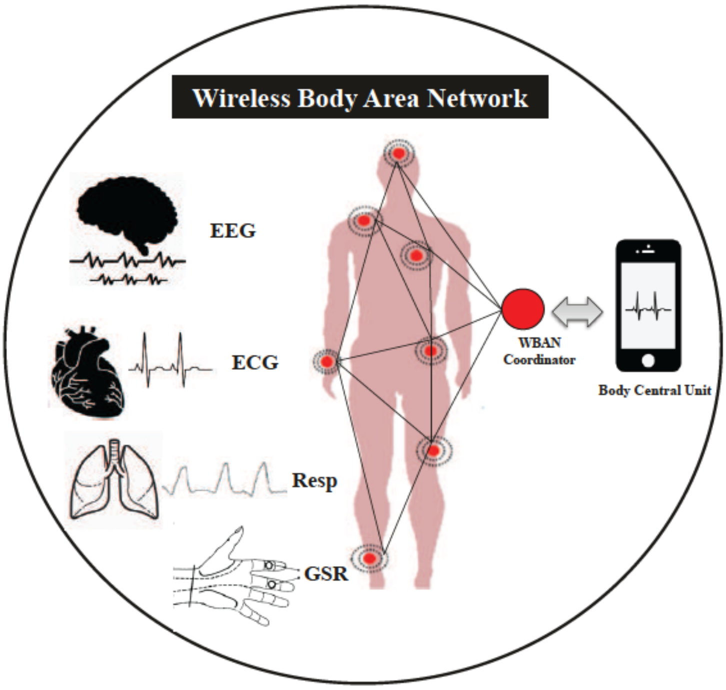


Figure 2.11: Wireless Body Area Network.

Table 2.2: Wireless biosensors and their characteristics

Biosensor	**Data Rate**	**Number of Nodes**	**Bit Error Rate**	**Desired Battery Life**
ECG	72 kB/s	< 6	$< 10^{-10}$	>one week
EEG	86.4 kB/s	< 6	$< 10^{-10}$	>one week
EMG	1.536 MB/s	< 6	$< 10^{-10}$	>one week
BVP	< 10 kB/s	< 10	$< 12^{-10}$	>one week
Temperature	< 10 kB/s	< 10	$< 12^{-10}$	>one week
Respiration	< 10 kB/s	< 10	$< 12^{-10}$	>one week
GSR	< 10 kB/s	< 10	$< 12^{-10}$	>one week

- MIT dataset collects data from drivers in order to measure their stress level and not in normal conditions.
- eENTERFACE acquires only brain signals via functional near-infrared spectroscopy (fNIRS), face video and scalp EEG signals, without considering pheripheral physiological signals.
- The MMI database includes images and videos captured from both frontal and profile view of 25 subjects.
- HUMAINE database varies in terms of size (8 to 125 participants) and modalities. It consists of three naturalistic and six induced reaction databases, among which it includes the Belfast database [32].

For the research purposes pursued in this work DEAP, MAHNOB-HCI and DECAF dataset have been studied in detail. **DEAP** is a database for emotion analysis using physiological signals acquired with Biosemi ActiveTwo system[3]: electroencephalogram (32 channels at 512Hz), skin conductance level, respiration amplitude, skin temperature, electrocardiogram, blood volume by plethysmograph, electromyograms of zygomaticus and trapezius muscles, electrooculogram, and face video (only for 22 participants). Participants watched 40 one-minute music videos, rating each video in terms of arousal, valence, and dominance, as well as like/dislike and familiarity with the videos. Stimulus selection was proposed using affective tags from the *last.fm* music website. Results showed negative correlations for arousal in the theta, alpha and gamma band. Valence showed the strongest correlations wiht EEG signal in all frequency bands. In theta and alpha frequencies, an increase of valence led to an increase of power. For the beta frequency band a central decrease was found. Increased beta power was associated with positive emotional self induction. A positive correlation of valence includes beta and gamma bands.

MAHNOB-HCI is a multimodal database for affect recognition and implicit tagging. Recorded signals are: 6 video tracks of face, speech from a head microphone, eye gaze, pupil size, electrocardiogram, electroencephalogram (32 channels), skin temperature and respiration amplitude (all biosignals at 256Hz). Two experiments were conducted. In the first test, participants watched 20 emotional videos and gave feedback in terms of arousal, valence, dominance and predictability. In the second experiment, 28 images and 14 video fragments were showed to participants accompanied by a word tag; agreement or disagreement with the displayed tags was assessed by participants. Authors proposed emotion recognition results from three modalities

[3]http://www.biosemi.com

Table 2.3: Emotional dataset review

Dataset	**N. subjects**	**Reaction: Posed vs Spontaneous**	**Stimuli: natural vs induced**	**Audio**	**Visual**	**Biosignals**	**Brain Signal**	**Eye gaze**
MIT [96]	17	spontaneous	natural	no	no	yes	no	no
eENTERFACE [114]	19	spontaneous	induced	no	yes	no	EEG	no
MMI [89]	16	both	induced	no	yes	no	no	no
HUMAINE [33]	multiple	spontaneous	both	yes	yes	yes	no	no
VAM [45]	19	spontaneous	natural	yes	yes	no	no	no
SEMAINE [81]	20	spontaneous	induced	yes	yes	no	no	no
EMDB [17]	32	spontaneous	induced	no	no	yes	no	no
DEAP [65]	32	spontaneous	induced	no	yes (22 subj.)	yes	EEG	no
MAHNOB-HCI [119]	27	spontaneous	induced	yes	yes	yes	EEG	yes
DECAF [1]	30	spontaneous	induced	yes	yes	yes	MEG	no

(peripheral physiological signals, EEG and eye gaze) and the results of the classification over the two best modalities were fused. Rusults showed classification on gaze data performed better than EEG and pheripheral signals. Classification using physiological signals alone gave the worst results. The fusion of eye gaze and EEG classification modalities outperformed all single modalities.

DECAF is a multimodal dataset for DECoding user physiological responses to AFfective multimedia content. It contains brain signals acquired using the Magnetoencephalogram (MEG) sensor and explicit/implicit emotional responses of 30 subjects while viewing 40 music videos (the same as in the DEAP dataset) and 36 movie clips. Experiments demonstrated emotions are more strongly and consistently evoked by movie stimuli. Acquired physiological data are: magnetoencephalogram, horizontal electrooculogram, electrocardiogram, electromyogram of the trapezius muscle and near-infrared face video. Subject's self-assessment is in terms of arousal, valence and dominance and it was compared with DEAP. Results demonstrated dominance may be hard to qualify in a movie, but it was found to be relevant with regard to musical compositions. Furthermore, using a linear SVM classifier the higher spatial resolution of MEG was shown to improve affect recognition with respect to EEG sensing. MEG signals effectively convey arousal and dominance, while peripheral signals efficiently represent valence. DECAF also contains continuous-time emotion annotations for movie clips by seven experts in order to evaluate dynamic emotion prediction.

2.2 Knowledge-based affective computing

In literature many classification techniques were tested for emotion recognition through physiological signal analysis. The recognition of a user's emotional state can be obtained also via a rule-based approach with predicate logic or other reasoning techiniques. A rule-based system consists of if-then rules, a collection of facts, and an interpreter controlling the application of the rules. A simple if-then rule in the form "if x is A, then y is B" consists of a *premise* or *antecedent* (the if-part), and a *consequent* or *conclusion* (the then-part). When the premise is known to hold in a scenario, the conclusion can be drawn. In [58] a rule based approach was adopted for multimodal emotion recognition. The framework is built around if-then rules using certainty factors to capture uncertainty of individual features. Authors in [91] developed a rule-based decision model for emotion recognition. They showed knowledge-based rules can be developed using human annotator agreement from temporal and 3D data from the visual channel; furthermore, emotion recognition accuracy of a system can be augmented by using the rule-based decision model in combination with supervised learning techniques. Nevertheless, the state of the art lacks to process raw sensor data and translate them in machine-understandable knowledge. Semantic Web technologies provide a reference representation formalisms for raw sensor data to make them machine-interpretable and to interlink data with existing resources on the Web.

2.2.1 Knowledge Representation

A Knowledge Representation System (KRS) is a software information system able to exploit information incapsulated in a Knowledge Base (KB) and derive implicit knowledge through automated reasoning procedures. KRS consist of five main components as depicted in Figure 2.12:

- *Terminological Box (TBox)*: a finite set of explicit definitions and it introduces the application domain vocabulary. It represents the conceptual model of a fragment of the world (*ontology*) and constitutes the intensional knowledge which does not usually change.
- *Assertion Box (ABox)*: contains statements relating to concrete instances of the world (individuals) expressed using the terms in the TBox. It represents the concrete model of a fragment of reality and it is the extensional knowledge which is subject to occasional or even constant change.
- *Reasoning services*: software procedures able to deduce new knowledge from the information explicitly asserted in the KB.
- *Access Interface*: an appropriate Application Programming Interface (API) allowing software applications to manipulate and query the knowledge contained in the KB.
- *Editing interface*: software application allowing the management of the TBox and ABox contents by a human operator.

The reference paradigm for Knowledge Representation is First Order Logic (FOL). FOL assumes that all representations concern a non-empty set of individuals, called *domain*. In particular, this thesis refers to a family of logic formalisms for KR in a subset of FOL called Description Logics (DLs) [4]. In DLs the basic syntax elements are:

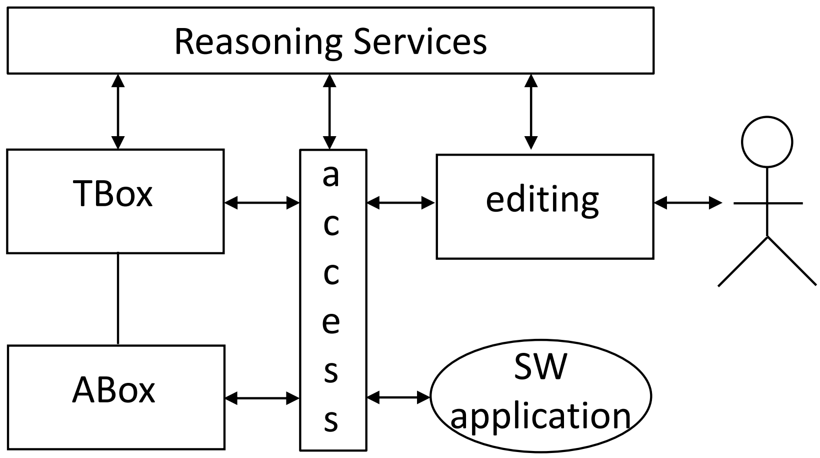


Figure 2.12: Knowledge Representation System.

- *concepts* or *classes*, representing sets of objects in the domain;
- *roles* or *properties*, which are binary relations between concepts;
- *individuals*, class instances representing objects in the domain of interest.

These elements can be combined by means of *constructors* to create DL *expressions*, specified in a formal semantics which associates an interpretation $\mathcal{I}$ to each term. Concept *conjunction* is interpreted as set intersection: $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$. Concept *disjunction* is interpreted as set union: $(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$. The connector $\neg$, if present, is the interpretation of the complement operator. More constructs exist, and each language in the DL family is characterized by the set of supported constructs, which determines both language expressivity and computational complexity of reasoning procedures. Constructs allow to create of semantic expressions:

- *definition*: a new concept can be defined in terms of other ones. For example, an emotion such as *Panic* can be defined as an emotion combining fast breathing, lower skin temperature causing shiver and pallor by writing this declaration:
  $Panic \equiv Fear \sqcap Becoming\_pale \sqcap Accelerated\_Breath \sqcap Shivering$
- *inclusion* (or *subsumption*): represent inclusion relations between arbitrary concept expressions of the form: $Panic \sqsubseteq Fear$

Also in this case the semantics of a definition is interpreted as a logical equivalence, while subsumption is interpreted, in turn, as set inclusion.

Table 2.4: Syntax and semantics of $\mathcal{ALN}$ constructors

Name	**Syntax**	**Semantics**
Top	$\top$	$\Delta^{\mathcal{I}}$
Bottom	$\bot$	$\emptyset$
Intersection	$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
Atomic negation	$\neg A$	$\Delta^{\mathcal{I}} \setminus A^{\mathcal{I}}$
Universal quantification	$\forall R.C$	$\{d_1 \mid \forall d_2 : (d_1, d_2) \in R^{\mathcal{I}} \rightarrow d_2 \in C^{\mathcal{I}}\}$
Number restriction	$\geq nR$	$\{d_1 \mid \sharp\{d_2 \mid (d_1, d_2) \in R^{\mathcal{I}}\} \geq n\}$
	$\leq nR$	$\{d_1 \mid \sharp\{d_2 \mid (d_1, d_2) \in R^{\mathcal{I}}\} \leq n\}$

This work will refer mainly to the *Attributive Language with unqualified Number restrictions* ($\mathcal{ALN}$) DL. Constructs of $\mathcal{ALN}$ are reported in what follows:

- $\top$, *universal concept.* All the objects in the domain.
- $\bot$, *bottom concept.* The empty set.
- A, *atomic concepts.* All the objects belonging to the set A.
- $\neg A$, *atomic negation.* All the objects not belonging to the set A.
- $C \sqcap D$, *intersection.* The objects belonging to both C and D.
- $\forall R.C$, *universal restriction.* All the objects participating in the R relation whose range are all the objects belonging to C.
- $\exists R$, *unqualified existential restriction.* There exists at least one object participating in the relation R.
- $(\geq nR)$[4], $(\leq nR)$, $(= nR)$[5], *unqualified number restrictions.* Respectively the minimum, the maximum and the exact number of objects participating in the relation R.

Table 2.4 lists the semantics of all the constructs used in $\mathcal{ALN}$ logic.

[4]Notice that $\exists R$ is equivalent to $(\geq 1R)$

[5]Notice that $(= nR)$ is a shortcut for $(\geq nR) \sqcap (\leq nR)$

Table 2.5: Correspondence between OWL and DL syntax

OWL syntax	**DL syntax**
owl : Thing	$\top$
owl : Nothing	$\bot$
owl : Classrdf : ID = "C"	C
owl : ObjectPropertyrdf : ID = "R"	R
rdfs : subClassOf	$\sqsubseteq$
owl : equivalentClass	$\equiv$
owl : disjointWith	$\neg$
owl : intersectionOf	$\sqcap$
owl : allValuesFrom	$\forall$
owl : someValuesFrom	$\exists$
owl : maxCardinality	$\leq$
owl : minCardinality	$\geq$
owl : cardinality	$=$

2.2.2 Web Ontology Language

OWL [126] (*Web Ontology Language*) is the knowledge representation languages for defining ontologies on the Semantic Web and annotating Web resources. The term OWL inglobes OWL 1 and OWL 2. OWL uses RDF/XML as main data model to represent classes, properties, individuals and data values. Three different levels of OWL sub-languages are defined, from the least to the most expressive:

- *OWL-Lite*. It allows very simple taxonomies and constraints on relation between classes.

- *OWL-DL*. It allows an extensive expressiveness while safeguarding computational decidability. The name recalls the direct correspondence between OWL DL and DLs.

- *OWL-Full*. It provides extreme flexibility of expression. The lack of restrictions is paid in terms of no computational decidability guarantee.

The subset of OWL-DL tags allowing to express the $\mathcal{ALN}$ DL is presented in Table 2.5.

The choice of OWL dialects is a tradeoff between expressiveness with its complexity and the availability of software support. This relationship is shown in Figure 2.13: as the language dialects become more expressive, they require more complex software to support them.

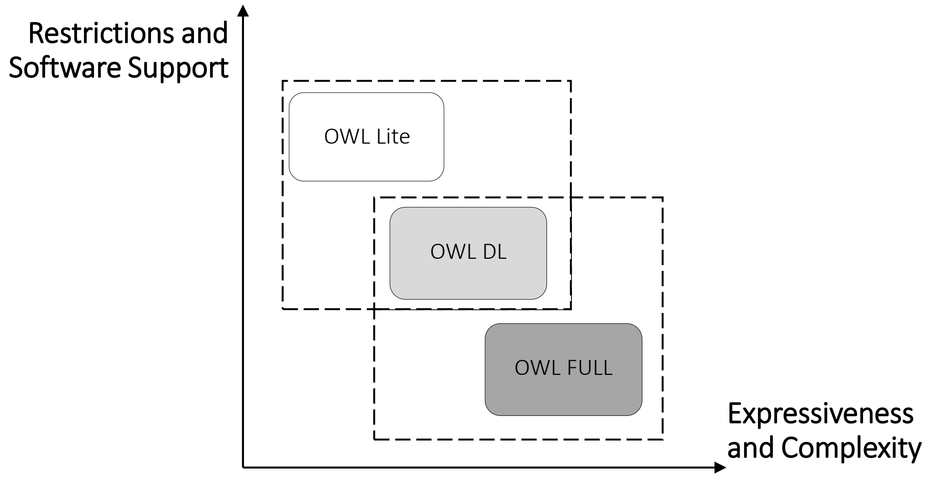


Figure 2.13: OWL dialects.

In OWL 2, there are three new profiles: *OWL 2 EL*, *OWL 2 QL* and *OWL 2 RL*. Each profile applies for a specific use case and offers a different tradeoff between expressiveness and reasoning efficiency. OWL Lite syntax is a subset of OWL DL, which is a subset of OWL Full, while OWL 2 profiles are independent among them. OWL 2 supports a variety of syntaxes to store ontologies and to exchange them among applications:

- *RDF/XML* is an XML syntax for Resource Description Framework (RDF); it is the only syntax that must be supported by all OWL 2 tools to guarantee interoperability.

- *OWL/XML* is an XML serialization for OWL 2 that mirrors its structural specification.

- *Functional Syntax* is a functional-style syntax whose main purpose is to specify the structure of the language.

- *Manchester Syntax* is a more "readable" compact syntax for OWL 2 ontologies.

- *Turtle* the Terse RDF Triple Language is a textual syntax for RDF allowing RDF graphs to be completely written in a compact and "natural" text form, with abbreviations for common usage patterns and datatypes.

2.2.3 Fuzzy Description Logics

Incompleteness and vagueness are intrinsic properties in various application domains, such as in the realm of emotions and their dynamic evolution. There are several approaches to the representation of uncertain knowledge in KR and particularly in Description Logics. Two of them are probabilistic logic [51] and fuzzy logic [127]. Fuzzy logic was proposed by Zadeh [128] to manage vague and imprecise knowledge.

Given a set of elements X, a fuzzy subset A of X is defined through a membership function $\mu_A(x)$, which assigns any $x \in X$ a value in the real number interval $\{0, 1\}$. In other words, elements can belong to a set with some *degree of truth*; particularly, 0 means no membership and 1 full membership. On the contrary, in classical set theory elements either belong to a set or not. Typical fuzzy membership functions are shown in Figure 2.14.

A main issue is the difficulty of constructing appropriate membership functions. One easy and often satisfactory method is to uniformly partition the range of values into 3 or 5 fuzzy sets using either trapezoidal or triangular functions.

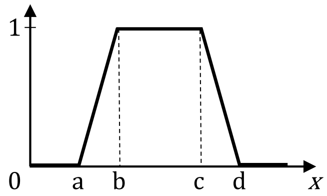
The goal of fuzzy Description Logics is to characterize notions which cannot be properly defined with a "crisp" numerical bound. Fuzzy DLs was introduced as an extension to classical DLs with the aim of dealing with *fuzzy/vague/imprecise information* (for more details see [11, 122]). In Fuzzy DLs, there are:
- *fuzzy concepts* representing classes of objects;
- *fuzzy roles* (a.k.a. *properties*) joining pairs of objects;
- *individuals* specific named objects;
- *fuzzy datatypes* (or fuzzy concrete domains) defined over an interval of the rational numbers.

In fuzzy DLs, an object may be an instance of a fuzzy concept to some degree in $[0, 1]$, while in the non-fuzzy case an object is either an instance of a concept or not. *Axioms* are statements which represent is-a relations between concepts. The logical statement has a *degree of truth* allowing to define new fuzzy concepts from other ones.

2.2.4 Affective computing ontologies

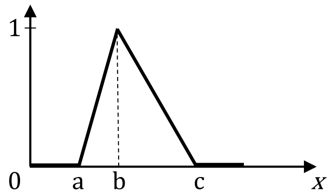
The suitable representation of emotions is an important issue in affective computing. In literature, different emotion ontologies have been developed to model the domain of human emotions and provide a vocabulary for semantic-based approaches. Mathieu presented an ontology to describe lexicon in the field of feelings and emotions, labelling words as positive, negative or neutral

Trapezoidal function *trz(a, b, c, d)*



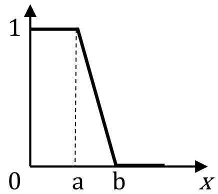
$$trz(x;\ a;\ b;\ c;\ d) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x \le b \\ 1, & b < x \le c \\ \dfrac{d-x}{d-c}, & c < x \le d \\ 0, & x > d \end{cases}$$

Triangular function *tri(a, b, c)*



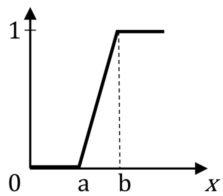
$$tri(x;\ a;\ b;\ c) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x \le b \\ \dfrac{c-x}{c-b}, & b < x \le c \\ 0, & x > c \end{cases}$$

Left shoulder function *ls(a, b)*



$$ls(x;\ a;\ b) = \begin{cases} 1, & x \le a \\ \dfrac{b-x}{b-a}, & a < x \le b \\ 0, & x > b \end{cases}$$

Right shoulder function *rs(a, b)*



$$rs(x;\ a;\ b) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x \le b \\ 1, & x > b \end{cases}$$

Figure 2.14: Typical fuzzy membership functions.

[79]. Francisco *et al.* modeled an emotional ontology based on DL for speech synthesis, where each emotional concept is defined in terms of a range of values along the three-dimensional space of emotions: this allows the system to make inferences in order to find a relationship between an arbitrary point in the dimensional space and a label in the categorical emotional space [40]. The Swiss Centre for Affective Sciences, in collaboration with the University at Buffalo, developed an *Emotion Ontology* to describe affective phenomena including their bearers, the different types of emotions, moods, etc., their different parts and dimensions of variation, facial and vocal expressions, and the role of emotions and affective phenomena in general in influencing human behavior [49]. Lopez *et al.* proposed an ontology for describing emotion detection and expression systems taking contextual and multimodal elements into account [75]. Similarly, Grassi developed an ontology for the annotation of emotions in multimedia data, which could be extended according to the user's purpose [44]. Authors in [48] found relationships between diseases and environmental causes (stress, family conditions, drugs, climate, pollution, noise), including even weather parameters and mood, *e.g.*, people are happier when it is sunny. Fang *et al.* trying to find the relationships between colors and the meridian system of traditional Chinese medicine building an ontology for clinical applications [38]. This focuses on the point of Chinese Medicine and energy medicine, and discusses the emotions and mind. Benta *et al.* proposed an ontology for context-aware applications that allows to express the complex relations among the affective states and between these and other context parameters [6]. *Onyx* is an ontology to represent the emotion analysis process and its results, as well as annotating lexical resources and connect results from different applications [113]. The *Emotion Ontology for Context Awareness* (*EmOCA*) describes and supports reasoning on emotional context in order to improve emotion detection based on bodily expressions, demonstrating how context influences human experiences [8]. *BIO_EMOTION* is another ontology-based context model for emotion recognition, describing user's context, profile and EEG data supporting inferences on the users' emotional states [130]. In [124] the *EmotionO+* ontology models physiological signals and a reasoning engine infers affective states providing an emotion context knowledge sharing platform for heterogeneous healtcare systems.

Even if not standardized into an ontology, the Humaine Project[6] and the W3C's Emotion Markup Language Incubator Group[7] defined languages for emotions description, EARL (Emotion Annotation and Representation Lan-

[6]http://emotion-research.net/

[7]https://www.w3.org/2005/Incubator/emotion/

guage) and EmotionML (Emotion Markup Language) respectively. The EARL language offers a powerful XML Schema for the annotation of audio-video databases. EmotionML is a general-purpose markup language for emotions and related states.

2.2.5 Inference services

Semantic annotations endowed with formal machine-understandable meaning provide structured information to reasoning algorithms. Semantic matchmaking is the process to find best matches between a request R and a resource description S, when both are satisfiable concept expressions w.r.t. a reference ontology. The assumptions that underpin semantic matchmaking are:

- *Open World Assumption.* The absence of a characteristic in a description should not be interpreted as a constraint of absence, but as unknown or irrelevant information.

- *Non-symmetric evaluation.* A matchmaking system may give different evaluations to the match between a resource S and a request R, depending on whether it is trying to match S with R, or R with S.

The inference services considered in the OWL 2 documentation[8] are labelled as "standard". Two of them are used for semantic matchmaking:

- *Concept Satisfiability.* Given an ontology $\mathcal{T}$ -modeling the domain of interest- and a resource description R referring to $\mathcal{T}$ R is satisfiable in $\mathcal{T}$ if at least one model of $\mathcal{T}$ exists assigning a non-empty extension to R. In formulae: $\mathcal{T} \models R \sqsubseteq \bot$.

- *Concept Subsumption.* Given an ontology $\mathcal{T}$ and two resource descriptions *R1, R2* both referring to $\mathcal{T}$ if *R2* is more general than *R1*, then *R1* is subsumed by *R2*. In formulae: $\mathcal{T} \models R1 \sqsubseteq R2$.

Other standard inference services that can be implemented through satisfiability and subsumption are: *consistency checking* and *classification*. The former ensures no contradiction of the facts contained in the ontology, the latter checks the subsumption relation between every concept in $\mathcal{T}$ in order to create a complete class hierarchy. Both Subsumption and Satisfiability are useful when an exact match is required between a user request and a set of resource profiles. Consequently, given a resource profile description S and a user preference request description R, all classic Semantic Web engines were designed to only distinguish among these match classes [25, 72]:

[8]http://www.w3.org/TR/owl2-semantics/#Inference_Problems

- *Exact*. All the features requested in R are exactly provided by S and vice versa. In formulae $\mathcal{T} \models R \Leftrightarrow S$.
- *Full-Subsumpion*. All the features requested in R are contained in S. In formulae: $\mathcal{T} \models R \Rightarrow S$.
- *Potential-Intersection*. The features offered in S are intersected with the ones requested in R. In formulae: $\mathcal{T} \not\models \neg(R \sqcap S) \sqsubseteq \bot$.
- *Partial-Disjoint*. There is a conflict between the features offered in S and the ones requested in R. In formulae: $\mathcal{T} \models \neg(R \sqcap S) \sqsubseteq \bot$.

In real-world application scenarios, featuring articulated descriptions from heterogeneous providers, this is often inadequate, as exact and full matches are rare while partial and potential ones are prevalent. Standard inferences do not allow comparing matches within the same class. On the contrary, non-standard inference services *Concept Contraction* and *Concept Abduction* [112] support approximate matches, semantic ranking and explanation of results.

Concept Contraction. If R and S are incompatible. In formulae: $\mathcal{T} \models R \sqcap S \sqsubseteq \bot$, the Concept Contraction Problem (CCP) consists in finding a pair of concepts (G,K) such that $\mathcal{T} \models R \equiv G \sqcap K$, and $\mathcal{T} \not\models K \sqcap S \sqsubseteq \bot$. G (*Give up*) is the explanation about what in R is incompatible with S, while K (*Keep*) represents the compatible part.

Concept Abduction. If there is compatibility between R and S but R does not match S fully, the Concept Abduction Problem (CAP) consists in finding a concept H such that $\mathcal{T} \models S \sqcap H \sqsubseteq R$. The concept expression H (*Hypothesis*) represents what should be hypothesized in S in order to completely satisfy R.

In both CCP and CAP, a penalty value is computed in order to evaluate the goodness of match approximation, which is the semantic distance of the description w.r.t. the request. These penalties can be combined through a *utility function*, such as:

$$d(S,R) = 100\left(1 - \frac{penalty_{(c)} + penalty_{(a)}}{max\ penalty_{(a)}}\right) \tag{2.1}$$

where $penalty_{(c)}$ and $penalty_{(a)}$ are the penalties induced by Concept Contraction and Concept Abduction between the service/resource annotation and the request. Penalty is normalized w.r.t. the maximum possible semantic distance from the request S, *i.e.*, the CAP penalty from the most generic DL concept (denoted $\top$); this distance depends only on the reference ontology and the request, not on individual resources. Finally, semantic affinity is

$\texttt{High\_Temperature} \sqsubseteq \texttt{Controlled\_Temperature}$
$\texttt{Medium\_Temperature} \sqsubseteq \texttt{Controlled\_Temperature}$
$\texttt{Low\_Temperature} \sqsubseteq \texttt{Controlled\_Temperature}$
$\texttt{Controlled\_Temperature} \sqsubseteq \texttt{Temperature}$
$\texttt{Room\_Temperature} \sqsubseteq \texttt{Temperature}$ $\texttt{High\_Humidity} \sqsubseteq \texttt{Controlled\_Humidity}$
$\texttt{Medium\_Humidity} \sqsubseteq \texttt{Controlled\_Humidity}$
$\texttt{Low\_Humidity} \sqsubseteq \texttt{Controlled\_Humidity}$
$\texttt{Controlled\_Humidity} \sqsubseteq \texttt{Humidity}$
$\texttt{Natural\_Humidity} \sqsubseteq \texttt{Humidity}$
$\texttt{High\_Oxygen} \sqsubseteq \texttt{Controlled\_Oxygen}$
$\texttt{Medium\_Oxygen} \sqsubseteq \texttt{Controlled\_Oxygen}$
$\texttt{Low\_Oxygen} \sqsubseteq \texttt{Controlled\_Oxygen}$
$\texttt{Controlled\_Oxygen} \sqsubseteq \texttt{Oxygen}$
$\texttt{Natural\_Oxygen} \sqsubseteq \texttt{Oxygen}$
$\texttt{Direct\_Lighting} \sqsubseteq \texttt{Lighting\_Source}$
$\texttt{Indirect\_Lighting} \sqsubseteq \texttt{Lighting\_Source}$
$\texttt{ISO\_Pallet\_Rack} \sqsubseteq \texttt{Pallet\_Rack}$
$\texttt{Pallet\_Rack} \sqsubseteq \texttt{Stocking\_Equipment}$
$\texttt{Plastic\_Shelving} \sqsubseteq \texttt{Shelving}$
$\texttt{Shelving} \sqsubseteq \texttt{Stocking\_Equipment}$
$\texttt{Stocking\_Equipment} \sqsubseteq \texttt{Equipment}$
$\texttt{Hydraulic\_Drill} \sqsubseteq \texttt{Transport\_Equipment}$
$\texttt{Transport\_Equipment} \sqsubseteq \texttt{Equipment}$
$\texttt{High\_Maturity} \sqsubseteq \texttt{Maturity}$
$\texttt{Low\_Maturity} \sqsubseteq \texttt{Maturity}$
$\texttt{High\_Fragrance} \sqsubseteq \texttt{Fragrance}$
$\texttt{Low\_Fragrance} \sqsubseteq \texttt{Fragrance}$
$\texttt{disj}(\texttt{High\_Oxygen}, \texttt{Medium\_Oxygen}, \texttt{Low\_Oxygen})$
$\texttt{disj}(\texttt{Controlled\_Temperature}, \texttt{Room\_Temperature})$
$\texttt{disj}(\texttt{Controlled\_Humidity}, \texttt{Natural\_Humidity})$
$\texttt{disj}(\texttt{Controlled\_Oxygen}, \texttt{Natural\_Oxygen})$
$\texttt{disj}(\texttt{Direct\_Lighting}, \texttt{Indirect\_Lighting})$
$\texttt{disj}(\texttt{Shelving}, \texttt{Pallet\_Rack})$ $\texttt{disj}(\texttt{High\_Maturity}, \texttt{Low\_Maturity})$
$\texttt{disj}(\texttt{High\_Fragrance}, \texttt{Low\_Fragrance})$
$\texttt{Ripe\_Product} \sqsubseteq \exists\texttt{Has\_Climateric\_Maturation\_Degree} \sqcap \forall\texttt{Has\_Climateric\_Maturation\_Degree}.\texttt{High\_Maturity}$
$\texttt{Unripe\_Product} \sqsubseteq \exists\texttt{Has\_Climateric\_Maturation\_Degree} \sqcap \forall\texttt{Has\_Climateric\_Maturation\_Degree}.\texttt{Low\_Maturity}$
$\texttt{disj}(\texttt{High\_Temperature}, \texttt{Medium\_Temperature}, \texttt{Low\_Temperature})$
$\texttt{disj}(\texttt{High\_Humidity}, \texttt{Medium\_Humidity}, \texttt{Low\_Humidity})$

Figure 2.15: Axioms in the food transport ontology used in the case study

expressed by a percentage: in this way several resources can be compared w.r.t. a request and the resource with highest rank can be selected by the requester.

Non-standard inference services for semantic matchmaking are implemented in the Mini-ME embedded reasoning engine [116], which is used in prototypes developed in this thesis. Mini-ME is implemented in Java and uses the OWL API [53] to parse and manipulate KB in OWL 2 syntaxes.

As a clarifying toy example, let us consider the following resources in the posture templates KB.
An Italian food-service company exploits the semantic-based supply chain management process to assist products delivery and shipping in different ports. Given a delivery request, a semantic-based product allocation process is performed receiving as input the following data:

- the set of *products* $P = \{p_1, p_2, \ldots p_n\}$ to be delivered along with their semantically annotated description (referring to the ontology $\mathcal{T}$ shown in Figure 2.15), required quantity q_i and destination;
- the set of available *ship holds* $S = \{s_1, s_2, \ldots s_m\}$ with the relative annotations and freight capacities;

Descriptions of some goods follow as an example:

Golden_Delicious: $Apple \sqcap \forall\ Has\_Colour.Yellow \sqcap \forall\ Has\_Quality.Ordinary\_Quality \sqcap \forall\ Storage\_Temperature.Cold\_Temperature \sqcap \forall\ Storage\_Humidity.High\_Humidity \sqcap \forall\ Storage\_Oxygen.Natural\_Oxygen \sqcap Unripe\_Product \sqcap \forall\ Has\_Storage\_Equipment.ISO\_Pallet\_Rack.$

Wheat: $Wheat \sqcap \forall\ Has\_Colour.Yellow \sqcap \forall\ Has\_Quality.Ordinary\_Quality \sqcap \forall\ Storage\_Temperature.Room\_Temperature \sqcap \forall\ Storage\_Humidity.Low\_Humidity \sqcap \forall\ Storage\_Oxygen.Natural\_Oxygen \sqcap \forall\ Storage\_Lighting\_Source.Indirect\_Light \sqcap \forall\ Has\_Storage\_Equipment.Bulk.$

Caravelle: $Pumpkin \sqcap \forall\ Has\_Colour.Orange \sqcap \exists\ Has\_Colour \sqcap \forall\ Has\_Quality.Ordinary\_Quality \sqcap \exists\ Has\_Quality \sqcap \forall\ Storage\_Temperature.Low\_Temperature \sqcap \exists\ Storage\_Temperature \sqcap \forall\ Storage\_Humidity.Medium\_Humidity \sqcap \exists\ Storage\_Humidity \sqcap \forall\ Storage\_Oxygen.Natural\_Oxygen \sqcap \exists\ Storage\_Oxygen \sqcap High\_Fragrance \sqcap \forall\ Has\_Storage\_Equipment.ISO\_Pallet\_Rack \sqcap \exists\ Has\_Storage\_Equipment.$

Navelina: $Orange \sqcap \forall\ Has\_Colour.Orange \sqcap \exists\ Has\_Colour \sqcap \forall\ Storage\_Temperature.Cold\_Temperature \sqcap \exists\ Storage\_Temperature \sqcap \forall\ Storage\_Humidity.High\_Humidity \sqcap \exists\ Storage\_Humidity \sqcap High\_Fragrance \sqcap \forall\ Has\_Storage\_Equipment.ISO\_Pallet\_Rack \sqcap \exists\ Has\_Storage\_Equipment.$

Cavendish: $Banana \sqcap \forall\ Has\_Colour.Yellow \sqcap \exists\ Has\_Colour \sqcap \forall\ Has\_Quality.Ordinary\_Quality \sqcap \exists\ Has\_Quality \sqcap \forall\ Storage\_Temperature.Low\_Temperature \sqcap \exists\ Storage\_Temperature \sqcap Medium\_Fragrance \sqcap \forall\ Storage\_Humidity.High\_Humidity \sqcap \exists\ Storage\_Humidity \sqcap \forall\ Storage\_Oxygen.Natural\_Oxygen \sqcap \exists\ Storage\_Oxygen \sqcap Ripe\_Product \sqcap \forall\ Has\_Storage\_Equipment.ISO\_Pallet\_Rack \sqcap \exists\ Has\_Storage\_Equipment.$

The mobile matchmaker embedded in the port operator device is used to identify groups of compatible products. Then ship holds should be filled with products that do not interfere each other causing a general quality loss. The following steps are executed:

1. the process starts with the allocation of a new empty cluster ($cargo_1$);
2. the first product **Navelina** is added to $cargo_1$;
3. the second product **Golden_Delicious** is compatible with the $cargo_1$ description and it is added to the group;
4. the subsequent good **Cavendish** is matched against to $cargo_1$ but a

semantic incompatibility is detected because the products are climacteric and so they can travel in the same cargo only if they are in the same ripening stage. In this case Cavendish is not added to the group;

5. **Caravelle** is also incompatible with $cargo_1$;
6. finally, **Wheat** is not compatible with $cargo_1$ due to conflicting storage requirements about temperature. Cold storage is required for fruit shipping whereas room temperature is suitable for transport of **Wheat**.

After the first loop, the following cluster is defined: $cargo_1$ = {Navelina, Golden_Delicious}. During the process, the system show possible incompatible features among products by means of the *Concept Contraction* service. The same process is repeated with remaining products to define other product clusters: $cargo_2$ = {Cavendish}; $cargo_3$ = {Caravelle}; $cargo_4$ = {Wheat}. Then, each cluster will be allocated on the more suitable hold. Let us consider the following hold descriptions:

Hold_1: $\forall$ *Storage_Temperature.Refrigeration* $\sqcap$ $\forall$ *Storage_Humidity.Controlled_Humidity* $\sqcap$ $\forall$ *Storage_Oxygen.Natural_Oxygen* $\sqcap$ $\forall$ *Storage_Equipment.ISO_Pallet_Rack* $\sqcap$ $\forall$ *Storage_Lighting_Source.Indirect_Lighting*.

Hold_2: $\forall$ *Storage_Temperature.Low_Temperature* $\sqcap$ $\forall$ *Storage_Humidity.Medium_Humidity* $\sqcap$ $\forall$ *Storage_Oxygen.Natural_Oxygen* $\sqcap$ $\forall$ *Storage_Equipment.ISO_Pallet_Rack* $\sqcap$ $\forall$ *Storage_Lighting_Source.Direct_Lighting* $\sqcap$ $\forall$ *Storage_Lighting_Type.Dim_Light* $\sqcap$.

Hold_3: $\forall$ *Storage_Temperature.Controlled_Temperature* $\sqcap$ $\exists$ *Storage_Temperature* $\sqcap$ $\forall$ *Storage_Humidity.Medium_Humidity* $\sqcap$ $\exists$ *Storage_Humidity* $\sqcap$ $\forall$ *Storage_Oxygen.Natural_Oxygen* $\sqcap$ $\exists$ *Storage_Oxygen* $\sqcap$ $\forall$ *Storage_Equipment.ISO_Pallet_Rack* $\sqcap$ $\exists$*Storage_Equipment* $\sqcap$ $\forall$ *Storage_Lighting_Source.Indirect_Lighting* $\sqcap$ $\exists$ *Storage_Lighting_Source*.

Hold_4: $\forall$ *Storage_Temperature.Room_Temperature* $\sqcap$ $\exists$ *Storage_Temperature* $\sqcap$ $\forall$ *Storage_Humidity.Low_Humidity* $\sqcap$ $\exists$ *Storage_Humidity* $\sqcap$ $\forall$ *Storage_Oxygen.Natural_Oxygen* $\sqcap$ $\exists$ *Storage_Oxygen* $\sqcap$ $\forall$ *Storage_Equipment.Bulk* $\sqcap$ $\exists$*Storage_Equipment* $\sqcap$ $\forall$ *Storage_Lighting_Source.Indirect_Lighting* $\sqcap$ $\exists$ *Storage_Lighting_Source*.

First of all, a compatibility check is performed between cargoes and available holds. For each cargo the semantic distance is evaluated and the results obtained are ranked. In this way each cargo will be allocated to the hold which features better satisfy transport requirements and maximizing the ship payload. According to the ranking results shown in Table 2.6, Results denote that, for example, $cargo_1$ is not compatible with $hold_4$, due to storage requirements not fulfilled by the bay of ship, particularly $hold_4$ is a bay suited to load bulk cargoes while $cargo_1$ requires refrigerator ship. An explanation of contrasting features obtained exploiting Concept Contraction inference

	$cargo_1$	$cargo_2$	$cargo_3$	$cargo_4$
$hold_1$	0.133	n.c.	n.c.	n.c.
$hold_2$	n.c.	0.279	0.03	n.c.
$hold_3$	n.c.	0.133	0	n.c.
$hold_4$	n.c.	n.c.	n.c.	0.019

Table 2.6: Ranking results between cargoes and holds

service:

$$\textbf{give up:} \; \forall\ Storage\_Equipment.ISO\_Pallet\_Rack \sqcap \\ \forall\ Storage\_Humidity.High\_Humidity \sqcap \\ \forall\ Storage\_Temperature.Cold\_Temperature$$

Finally, the cargoes will be arranged on the holds as follows: $hold_1$ {Navelina, Golden_Delicious}; $hold_2$ {Cavendish}; $hold_3$ {Caravelle}; $hold_4$ {Wheat}.

2.3 State of the art: issues and limitations

The literature about emotion recognition through biosensors shows a standardized procedure to build EmI systems, summarized in the four following stages [86, 98]:

1. signal acquisition;
2. preprocessing,
3. feature extraction;
4. machine learning based classification.

Conventional methods, however, lack deeper insight into the suitability of the overall data mining techniques for handling the intrinsic hidden information encapsulated in the raw sensor data. Castellano *et al.* [18] developed a system to recognize emotional states via a multimodal approach, using non-verbal behaviour modalities: facial expression, body movement, gestures and speech are exploited to recognize eight emotions. The multimodal approach was shown to improve recognition accuracy by 10% with respect to the most successful unimodal system. The major drawback in machine learning based classification problem is to reduce the dimensionality of the feature space to overcome the risk of overfitting, and the classifier is usually obtained by offline analysis of data records from the subject. The majority of existing approaches use (i) relatively obtrusive sensing devices connected to

standard computer workstations via wired links and (ii) classical machine learning techniques for emotion classification from features extracted from physiological data series. As such, they are suitable only for laboratory settings and not easily applicable to real-world scenarios. Furthermore, machine learning outputs consist in opaque labels, lacking a meaningful and machine-understandable interpretation of detected information. This is a limit for possible exploitation of Affective Computing (AC) frameworks in many application scenarios.
Recent developments in Body Area Network (BAN) technologies allow local processing of data gathered from wearable sensors routed through multi-hop wireless links toward a portable computing device (*e.g.*, a smartphone) [21]. In BAN-based mobile real-time emotion detection systems, performance of the processing pipeline is critical in terms of both computational efficiency and classification accuracy. Furthermore, classification should return trivial labels, without a formally structured description about the characteristics of the elicited user emotion. For example, systems in [3, 43] incorporate various aspect of emotional system, but they are not available as open source and are not fully oriented towards the use of standard representations. The Semantic Web initiative generated standard logic-based languages for the representation of ontologies to enable machine-understandable characterization of knowledge domains. Emotion recognition approaches exploiting Semantic Web technologies appear as a promising research direction; actually, some proposals exist in literature, although they are a minority. Zhang *et al.* [131] proposed a system based on reasoning rules applying a Decision Tree, but mining was exploited only to map data to a single class. Furthermore, a rule-based system is useful only if there is an exact match between rules and the data: this is quite rare in complex domains like emotion characterization.

The complexity of emotions, the non-linearity of biosignals, the impossibility to find a single model to represent emotions can be faced by adopting fuzzy systems. They are generally robust and have the ability to process inaccurate and vague data. Mandrik *et al.* [77] combined the dimensional and categorical model using fuzzy logic. Two fuzzy systems in cascade were involved in the process. The first system computes arousal and valence values from physiological signals, then these values are used to generate emotional labels. The possibility of describing a label as an interval in a dimensional model was also investigated in [8].
The human emotional status cannot be directly measured through a single sensor. In [29] the emotional response was characterized as a collection of physiological changes brought about by certain brain regions. Nevertheless, emotions can be correlated to external and/or internal factors, which can be analyzed. The internal factors can be measured from different body parts

through biosensors.
Emotion ontologies [44, 75, 130] have been developed by to model the domain of human emotions and provide a vocabulary for semantic-based recognitions. These earlier efforts inspired the reference ontology modeled for this work. The most common entities will include the different emotional states, user's profile and context, as well as information extracted from physiological signals. Despite a common concepts, emotion ontologies do not involve fuzzy concepts. The proposal KB is a fuzzy OWL 2 KB of emotions providing terminology to describe a large variety of heterogeneous and vague physiological sensations through a formal conceptualization of entities and relationships between them.
Although a large amount of work has been carried out about fuzzy logic-based machine learning [22], the application of ML techniques to OWL 2 has been only moderately addressed (see [107] for a recent overview). The focus has been mostly on the problem of learning concept expressions without involving so-called numerical datatype restrictions. They, however, play an important role in emotion classification, because the available data gathered from sensors is numeric and sequential in time. Furthermore, emotion categorization models, created in psychology and AC research, are characterized by an inherent vagueness of terms and relationships. Expressing emotions, in fact, is affected by many subjective factors, including age, language and culture. This leads to difficulty in describing emotions objectively and accurately. New learning methodologies are thus needed, capable of dealing with both numerical input data and imprecise concepts. Fuzzy logics provide a suitable theoretical framework for such kind of problems, also useful to improve the human readability (see *e.g.*, [122, 123]). Fuzzy ontology-based machine learning techniques have been scarcely investigated in general [122] and not at all in the context of EmI. Many inductive algorithms can reach a good effectiveness, but can also produce rules that are hard to be understood for a human being. The readability, besides effectiveness, is also important. In order to achieve better human readability of the induced rules, fuzziness was adopted to describe emotional concepts. Together with the adoption of non-standard inference services supporting approximate matches, this appears to be an interesting ingredient to improve EmI system performance in terms of fine-graned emotion categorization, flexible classification and logic-based explanation of the outcomes.

Chapter 3

Emotion recognition through fuzzy DL learning and semantic matchmaking

This chapter describes the proposed framework and the prototype tool developed for its validation, first detailing the characteristics of this architecture and every single module that constitutes it.

3.1 Knowledge-based framework

The devised framework extends the previous emotion recognition works discussed in Chapter 2. Most existing approaches to detect emotions [99] preprocess signals provided by body sensors in order to remove noise and artifacts, extract meaningful features, select the most relevant ones and remove redundant or irrelevant ones to decrease the computational cost and avoid classifier over-fitting. The final classification step uses a machine learning technique, *e.g.*, k-Nearest Neighbor (k-NN), Regression Trees (RT), Bayesian Networks (BNT), Support Vector Machines (SVM), Artificial Neural Networks (ANN), Linear Discriminant Analysis (LDA).
In the proposed approach, workflow steps remain basically the same, but semantic-based enhancements change the way each step is performed. The overall framework aims at an EmI system able to recognize emotions and - depending on the application scenario - provide feedback impulses to the user.

3.1.1 Architecture

The overall architecture is depicted in Figure 3.1, showing the proposed system functionality and components.

A. Body Area Network. Autonomous sensor nodes can form a body area network (BAN) or body sensor network (BSN) [70]. Gathered physiological parameters are routed through multiple wireless links in the BAN to a portable device (*e.g.*, a smartphone) with constrained computational resources for real-time DAQ (Data Acquisition).

B. Raw data. In order to make emotion recognition accurate and reliable, the proposed system takes as input a combination of bio-signals associated with emotions experienced by the user. Physiological signals in response to stimuli are collected and used as the EmI system input. Before these data streams are fed into the computational model, feature extraction techniques are exploited.

C. Semantic annotator. The goal of this component is to build a semantic annotation combining physiological features and user contextual information.

D. Reference Knowledge Base. The semantic description is compared with annotations contained in a reference Knowledge Base. In the proposed architecture, the KB plays the role of *model* in typical machine learning approaches. It is created in a preliminary training step from a reference biosignal dataset.

E. Matchmaker. Non-standard inference services for semantic matchmaking described in Section 2.2.5, [112] produce the semantic description of the detected emotion. They are implemented in the Mini-ME mobile and embedded matchmaking and reasoning engine [116], designed to provide high performance on moderately expressive DLs, such as $\mathcal{ALN}$.

F. Biofeedback. The semantically annotated emotion description captures the best action to enhance user's affective states, giving a feedback impulse. The goal of biofeedback [80] is to make the subject aware of a problem in his/her affective state or behavior and stimulate a correction. Depending on the particular application scenario, emotional biofeedback can be electromechanical (through wearable actuator devices in the BAN), acoustic or visual.

Hereinafter, the five sequential steps (Figure 3.2) listed in the framework description are detailed.

- *Biosignal acquisition*: the freely accessible MAHNOB-HCI [119] and DEAP [65] databases for affect recognition and tagging were firstly used as ground truth.

- *Data preprocessing*: it is a crucial stage because raw data are affected

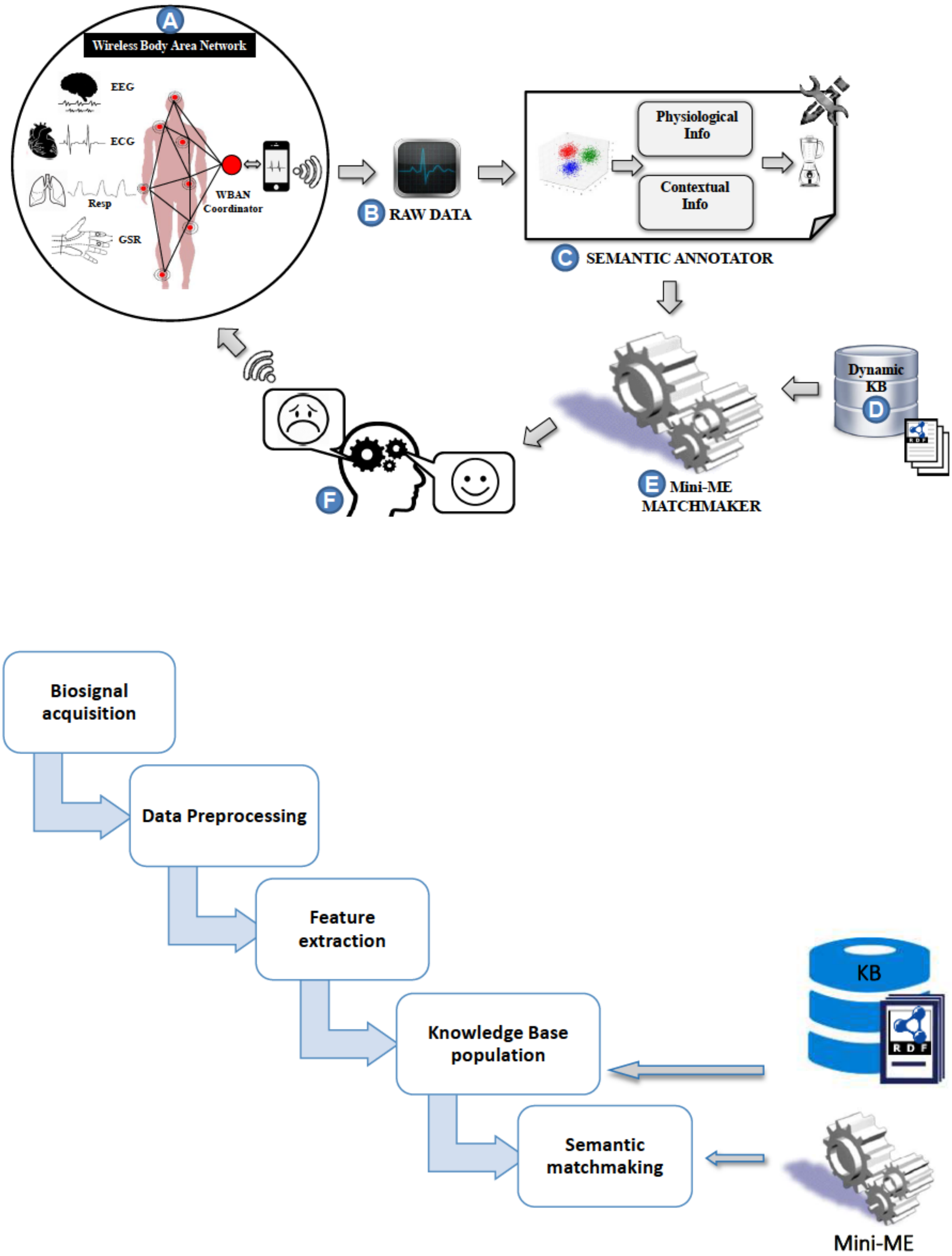


Figure 3.2: Processing steps.

by artifacts due to the limitations of sensing devices or to subject's involuntary movements (*e.g.*, blinking). This phase includes:

   a. resampling;
   b. removal of EOG artifacts;
   c. application of a frequency filter to eliminate the power frequency (50 Hz noise);
   d. normalization;
   e. signal segmentation in order to consider only the signal part with higher information content, excluding the initial and the final part.

- *Feature extraction and selection*: many physiological parameters can be derived from biosignals. Feature selection in the proposed framework was based on a survey [66] indicating the most significant biosignal features for emotion detection. The survey reviewed 134 publications of experimental investigations of emotional effects on peripheral physiological parameters. Further details are discussed in the Section 3.1.2.

- *Dynamic KB construction*: it is completed in two distinct phases. The TBox was modeled manually, considering for every emotion all the features in order to maps extracted features to the categorical model of emotion as shown in [66]. The ABox is automatically built annotating the features extracted from the biosignals of the reference dataset.

- *Semantic Matchmaking*: it allows to extract the emotion felt by the subject under examination through the application of non-standard inference algorithms.

A semantic-based approach allows a more effective information modeling describing the complex domain of emotions and affective states, so overcoming:

- issues related to the syntactic representation used by many current emotion recognition approaches;

- limitations of common machine learning techniques providing only a simplistic classification of emotion and producing opaque labels without capturing the relationships between biosignals and emotions.

3.1.2 Biosignal feature extraction

Biosignals must be preprocessed before the feature extraction phase. Preliminarly, all artifacts and outliers are eliminated. At least two noise types exist: environmental noise produced by electrical recording equipment, *i.e.*, the 50Hz power line noise, and physiological noise due to body movements such as muscle contraction. The second stage involves data reduction because a high number of data samples per subject observation (i.e., dependent variables) is produced, which may contain a large amount of noise leading to a worse separation of the various classes of interest. The baseline values are calculated and subtracted, in order to consider only the most relevant information part.
More specifically, in the proposed framework raw data was preprocessed as described above, resampled and EOG artifacts were removed. Before feature extraction, the pre- and post-trial baseline (measured before and after the trigger stimulus), in order to work only with the significant information content of the signal, and the line frequency by applying a low-pass filter with a cutoff frequency of 50 Hz. Furthermore, the ECG signal underwent a pre-processing step in order to allow its use. It was measured on 3 channels, EXG1 (top right of the chest, below the collarbone), EXG2 (top left of the chest, below the collarbone) and EXG3 (left leg). In order to have the complete cardiac track, lead II is calculated from a combination of the other two leads by applying Einthoven's Law as depicted in Figure 3.3. Positioning the electrodes on the limbs as in the picture, lead I records the voltage between the right and left arm, lead II the voltage between the right arm and the left leg, and lead III the voltage between the left arm and the left leg [14]. Einthoven's Law then follows immediately from Kirchhoff's Voltage Law:

$$LeadIII = LeadII - LeadI \tag{3.1}$$

The following software libraries were used to preprocess data:

- *EEGLAB*[1], a flexible MATLAB toolbox for processing continuous and event-related EEG, MEG and other electrophysiological data. It allows independent component analysis, time/frequency analysis, artifact rejection, event-related statistics, and several useful visualization modes.
- *Ledalab*[2], a MATLAB software for the analysis of skin conductance data (*i.e.*, EDA, GSR). It enables feature extraction via continuous and discrete decomposition analysis.

[1]EEGLAB: http://sccn.ucsd.edu/eeglab/
[2]Ledalab: http://www.ledalab.de/

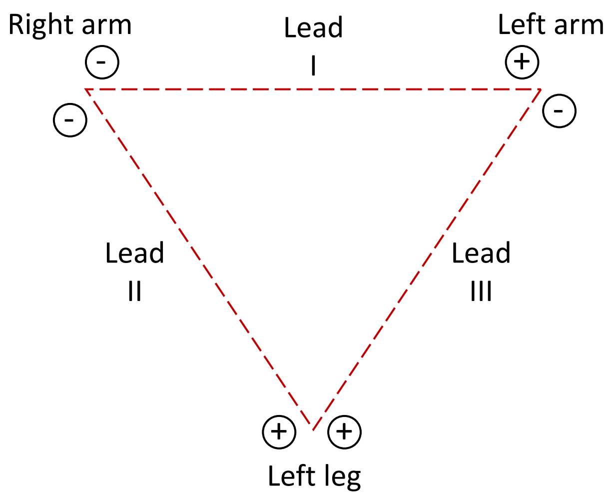


Figure 3.3: Einthoven's triangle.

- *Kubios*[3], a software tool for medical signal analysis, particularly for heart rate variability analysis.
- *BioSig*[4], an open source software library for biomedical signal processing. BioSig provides solutions for data acquisition, artifact processing, quality control, feature extraction, classification, modeling and data visualization.
- *libRASCH*[5], a library for risk stratification on the basis of non-invasive biosignals, such as ECGs, arterial pressure curves and respiration signals. Functionality and compatibility with multiple file formats are provided plugins.

After signal preprocessing features extraction was performed in MATLAB with BioSig and EEGLAB libraries. The six statistical features computed for each of the signals were obtained through Equation (3.2) to Equation (3.7). X_n is the value of the n^{th} sample of the raw signal, where $n = 1, ..., N$; $\tilde{\mathrm{X}}_n$ refers to the Z-score normalized signal: $(\mathrm{X}_n$-$\mu_x)/\sigma_x$.

1. average

$$\mu_x = \frac{1}{N}\sum_{n=1}^{N} X_n \tag{3.2}$$

[3]Kubios: http://www.kubios.com/
[4]BioSig: http://biosig.sourceforge.net/
[5]libRASCH: http://www.librasch.org/librasch/

2. standard deviation

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(X_n - \mu_x)^2} \tag{3.3}$$

3. average of the absolute value of the first differences

$$\delta_x = \frac{1}{N-1}\sum_{n=1}^{N-1}|X_{n+1} - X_n| \tag{3.4}$$

4. average of the absolute value of the first differences of the normalized signal

$$\tilde{\delta_x} = \frac{1}{N-1}\sum_{n=1}^{N-1}|\tilde{X}_{n+1} - \tilde{X}_n| = \frac{\delta_x}{\sigma_x} \tag{3.5}$$

5. average of the absolute values of the second differences

$$\gamma_x = \frac{1}{N-2}\sum_{n=1}^{N-2}|X_{n+2} - X_n| \tag{3.6}$$

6. average of the absolute values of the second differences of the normalized signal

$$\tilde{\gamma_x} = \frac{1}{N-2}\sum_{n=1}^{N-2}|\tilde{X}_{n+2} - \tilde{X}_n| = \frac{\gamma_x}{\sigma_x} \tag{3.7}$$

The above statistical features were chosen because they can be computed easily online, which makes them convenient for real-time recognition systems. The statistical features, however, do not exploit knowledge about the physical sources of the signals. Therefore further phisiological features were extracted. The choice of the most significant features to be extracted is based on a review of 134 experimental investigations of emotional affects on the autonomic Nervous System (ANS) response, which considered different physiological signals [66]. The features characterizing each involved signal type, along with a brief description of the variability range, are shown as follows:

1. Electroencephalography: alpha, beta, theta and delta frequency bands power, with ranges 13-15, 7.5-13, 2.5-8, and < 4 Hz respectively, were taken into account as EEG features [13].

2. Heart Rate (HR): three heart rate ranges are identified and categorized: low (< 60 bpm), normal ($60 \div 100$ bpm) and high (> 100 bpm) [78].

3. Beta/alpha ratio: it is extracted from EEG signal recording and indicates how relaxed a person is (and therefore the level of arousal), because beta waves are connected to an alert state of mind, whereas alpha waves are more dominant in a relaxed subject [13].

4. Heart Rate Variability (HRV): was categorized as high, low and very low with ranges 0.15-0.4, 0.04-0.15 and 0.003-0.04 Hz respectively. HRV is known to decrease with anxiety and increase with amusement [46].

5. Stroke Volume (SV): the normal range of stroke volume is from 0 ml to 250 ml. SV remains almost static for positive emotions, while it responds actively to negative emotions *e.g.*, it decreases with disgust, fear and sadness [3]. Low range is from 100 to 144 ml, normal from 10 to 250 ml, and high from 240 to 400 ml.

6. Cardiac Output (CO): can be derived from heart rate and stroke volume with the formula $HR \cdot SV/1000$[6]; its normal range is 4.0 to 8.0 L/min, low is < 4.0 L/min and high is > 8.0 L/min.

7. Systolic Blood Pressure (SBP): is labelled as low if < 90 mm Hg, normal from 90 to 140 mm Hg, and high if > 140 mm Hg. SBP increases with fear and anxiety.

8. Diastolic Blood Pressure (DBP): is labelled as low if < 60 mm Hg, normal from 60 to 90 mm Hg and high if > 900 mm Hg. DBP increases with anger, anxiety, and disgust, while it decreases with acute sadness [66].

9. Skin Conductance Response (SCR): it is considered low between 0 to 0.2 ms, normal from 0.1 to 1 ms and high from 0.85 to 1.5 ms [117].

10. Tidal Volume (Vt): The normal volume of air displaced between normal inspiration and expiration whitout applying an extra effort is represented as rapid breath from 100 to 150 ml/breath, quiet breath from 200 to 750 ml/breath and deep breath from 600 to 1200 ml/breath [15].

11. Oscillatory Resistance (Ros): three labels are adopted namely low from 0 to 0.49, normal from 0.4 to 0.88 and high from 0.5 to 1 [62].

12. Respiration Rate (RR): it is divided in low from 5 to 10, normal from 7 to 23 and high from 15 to 24 breath/min [62].

[6] Normal Hemodynamic Parameters and Laboratory Values. http://ht.edwards.com/scin/edwards/sitecollectionimages/edwards/products/presep/ar04313hemodynpocketcard.pdf

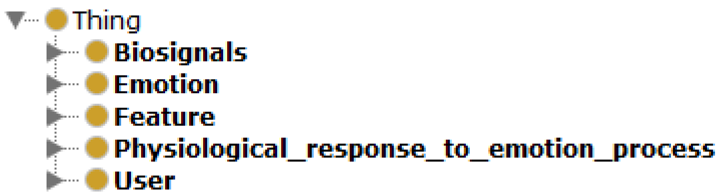


Figure 3.4: High-level ontology classes.

13. Nonspecific Skin Conductance Response (nSCR): implemented with three frequency categories, low from 0 to 2, normal from 1 to 3 and high from 2 to 5 per/min [62].

14. Skin Conductance Level (SCL): an indication of physiological arousal is given by SCL, which was divided in three ranges low from 0 to 2, normal from 2 to 25 and high from 20 to 25 mS [76].

15. Finger Temperature (FT): it was categorized as low from 65 to 75°F, normal from 75 to 85°F and high from 80 to 90°F [62].

As the reader may notice, value ranges for a certain feature were generally as overlapping, not disjoint. This was a deliberate choice, in order to take into account margins of subjective variability of physiological parameters. The complete subject description will be formed by the conjunction of concepts annotating each feature. Matchmaking will be able to detect the globally closest emotional state of the subject w.r.t. the ones derived from the training dataset by reasoning on such descriptions. The ontology modeled to allow feature labeling is described in Section 3.1.3, while the annotation process is explained in Section 3.1.4.

3.1.3 Knowledge base population

In order to support emotion recognition based on semantic matchmaking, an Affective Computing Ontology was modeled. It standardizes the main models for the representation of human emotions and their context. The language used is the OWL 2 fragment corresponding to the $\mathcal{ALN}$ DL. The *Protégé* editor was used to create it. Figure 3.4 shows the upper classes of the taxonomy: they are expandend in subsequent figures as follows:

- *Biosignal* (Figure 3.5): it represents the signals originating from the central nervous system such as EEG, EKG, electrooculogram, etc.

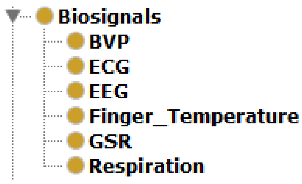


Figure 3.5: Subtree of Biosignal class.

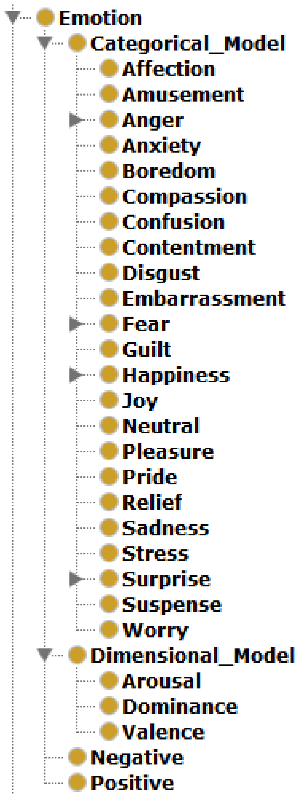


Figure 3.6: Subtree of Emotion class.

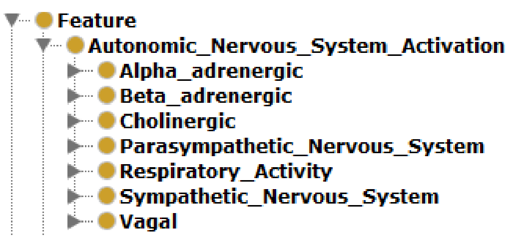


Figure 3.7: Subtree of Autonomic_Nervous_System_Activation class.

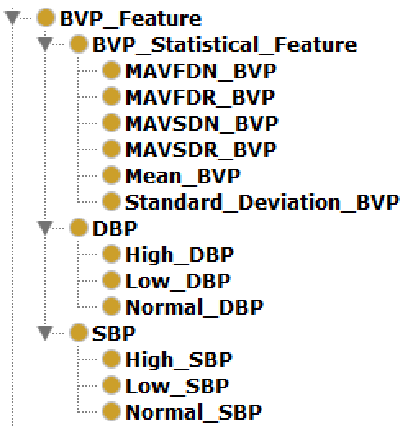


Figure 3.8: Subtree of BVP_Feature class.

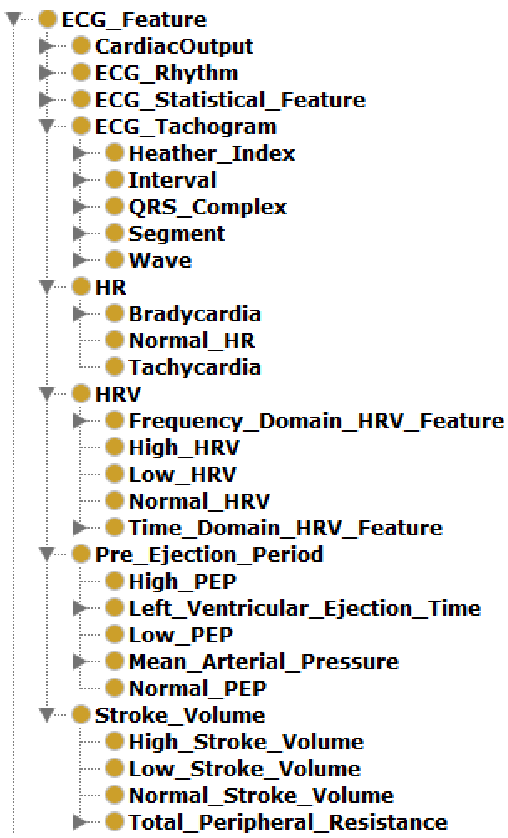


Figure 3.9: Subtree of ECG_Feature class.

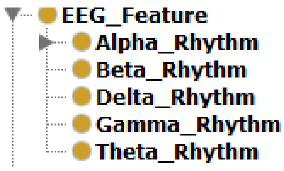


Figure 3.10: Subtree of EEG_Feature class.

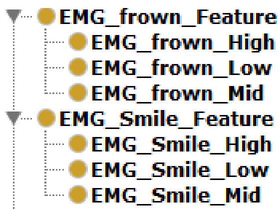


Figure 3.11: Subtrees of EMG_frown_Feature and EMG_Smile_Feature classes.

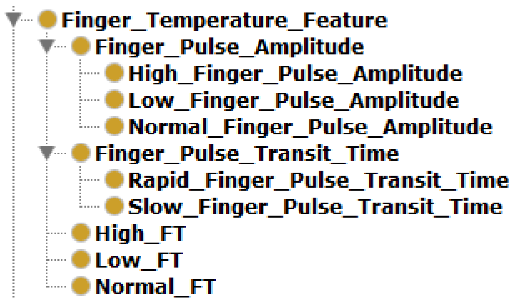


Figure 3.12: Subtree of Finger_Temperature_Feature class.

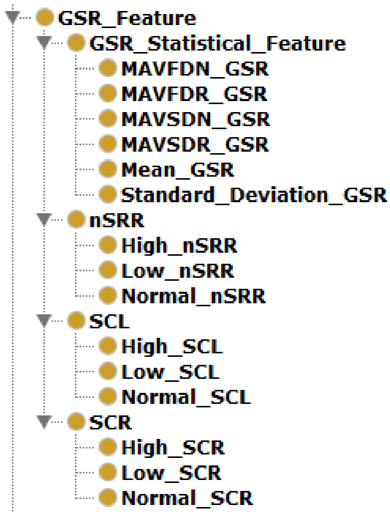


Figure 3.13: Subtree of GSR_Feature class.

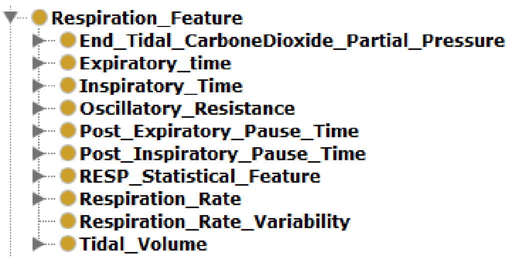


Figure 3.14: Subtree of Respiration_Feature class.

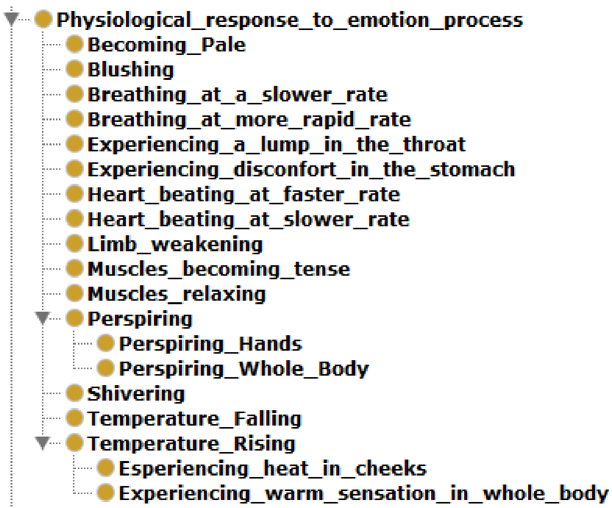


Figure 3.15: Subtree of Physiological_response_to_emotion_process class.

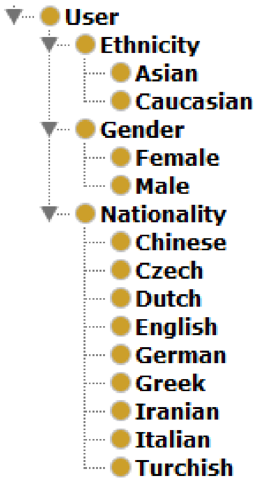


Figure 3.16: Subtree of User class.

- *Emotion* (Figure 3.6): defines the user's emotional state labeled according to the categorical and dimensional approaches. As explained in Section 2.1.2, the former is a discrete model, more suitable to determine an exact match with a specific emotional state; the latter is a continuous model, ideal for real-time monitoring of the user's emotional state. The class *Dimensional Model* is a linear combination of two or three dimensions of Valence-Arousal-Dominance (VAD) space.

- *Feature*: it models the features of each biosignal. Hereafter the subtrees of each category of features are shown:

  - *Autonomic_Nervous_System_Activation* (Figure 3.7): the autonomic activity is emotion-specific. Each component is modeled through three subclasses indicating the activation degree distinguished in low, medium and high. The *Alpha_adrenergic* receptors have a significant relevance in the control of the cardiovascular system functions, especially in high-arousal emotions, while the *Beta_adrenergic* ones are more involved in muscle relaxation conditions. The *Cholinergic* receptors, together with the *Parasympathetic nervous system*, allows modeling muscle contraction conditions, for example due to a fear emotion. The *Sympathetic nervous system* provides information about "fight or flight" conditions. *Respiratory activity* and *vagal* systems complete the physiological picture of an emotionally involved subject.
  - *BVP_Feature* (Figure 3.8): the blood volume pulse feature is sta-

tistically modelled, through mean, standard deviation and average of the first and second derivatives. Overall arterial pressure varies between the highest level of pressure seen at systole (systolic blood pressure or SBP) and the lowest level seen in diastole (the diastolic blood pressure or DBP), for these reasons diastolic and systolic levels are also taken into account.

- *ECG_Feature* (Figure 3.9): features extracted from ECG are the result of autonomic nervous system regulatory activities. This subtree models statistical features as well as the most common ECG waveform features used in cardiology. Further relevant features are the heart rate (to distinguish between bradycardia, *i.e.*, low rate and tachycardia, i.e., high rate)) and heart rate variability (a general term used to refer to changes in heart rate). Such features have been related to individual differences in attention, cognition and emotion in both child and adult populations. Sympathetic cardiac control has also been evaluated by measuring the duration of pre-ejection period (PEP, that is the period between the ventricular contraction and the start of blood ejection into the aorta) and the stroke volume, or the volume of blood pumped from the left ventricle with each beat of the heart.
- *EEG_Feature* (Figure 3.10): EEG activity (*i.e.*, increases/decreases) allows to distinguish between functional inhibitory and excitatory activities.
- *EMG_frown_Feature/EMG_Smile_Feature* (Figure 3.11): facial expressions of emotion, such as frowns, grimaces, smiles, are signs of emotional engagement, and are measured by EMG of facial muscles.
- *Finger_Temperature_Feature* (Figure 3.12): finger temperature is basically modeled as high, medium or low. Moreover, the *Finger_Pulse_Amplitude* and the *Finger_Pulse_Transit_Time* subclasses describe the reaction time to the temperature increases, in terms of amplitude and duration.
- *GSR_Feature* (Figure 3.13): GSR has been one of the most widely used response systems because it provides valuable clues to attention and emotion. The most useful electrodermal features in the context of continuous stimuli are SCL, SCR and frequency of NS-SCRs, because they can be measured on an ongoing basis over relatively long periods of time.
- *Respiration_Feature* (Figure 3.14): this class contributes to a com-

plete modeling of the various stages in the respiratory activity. It starts from the respiration rate, its variability and the oscillatory resistance, up to the time of expiration and inspiration distinct as fast and slow, and the time that elapses after each post-respiratory phase. The level of oxygen saturation and the amount of air ingested during a breath are also modeled.

- *Physiological_response_to_emotion_process* (Figure 3.15): this class models the physiological responses to emotion. For example in response to some kind of fear phenomena as *Becoming_Pale, Muscle_becoming_tense, Temperature_Falling* often occur, while the physiological responses of *Blushing, Heart_Beating_at_faster_rate, Temperature_Rising* are usually associated with a happiness emotion.

- *User* (Figure 3.16): models the user by characterizing its ethnicity, gender and nationality in order to assess how differences of such contextual information can influence the expression of emotions.

3.1.4 Clustering and annotation

The next step is semantic emotion annotation with respect to the reference ontology. This expresses in a formal manner the information associated with a subject using concepts, roles and axioms in the ontology in order to provide an unambiguous semantics. The annotation module was developed in Eclipse environment and written in the Java language with the OWL API support [53]. A semantic description is the conjunction of concepts representing user physiological information. Three main steps are performed in cascade for the construction of a semantic description:

1. features are extracted from the biosignal waveform;

2. each feature value is labelled with the corresponding conjunct concept from the ontology;

3. the complete semantic annotation is composed as the conjunction of the concepts.

The individual's description is formed by the conjunction of the user's physiological characteristics represented by the automatically generated labels. Each conjunct is generated by evaluating the value of the related feature. Some examples of the variability range for particular features with their respective semantic annotations are reported in Table 3.1.

Table 3.1: Examples of feature ranges and semantic annotation composition

Feature	**Range**	**Label**	**Subject value**	**Semantic annotation**
HR	< 60 $60{-}100$ > 100	Bradycardia Normal_HR Tachycardia	150	$\forall hasHR.Tachycardia$
Ros	$0 - 0.49$ $0.4 - 0.88$ $0.5 - 1.0$	Low_Ros Normal_Ros High_Ros	0.32	$\forall hasRos.Low\_Ros$
SV	$10 - 144$ $145 - 250$ $260 - 400$	Low_SV Normal_SV High_SV	80	$\forall hasSV.Low\_SV$
CO	< 4 $4.0 - 8.0$ > 8	Low_CO Normal_CO High_CO	3.5	$\forall hasCO.Low\_CO$
DBP	< 60 $60 - 90$ > 90	Low_DBP Normal_DBP High_DBP	95	$\forall hasDBP.High\_DBP$
SBP	< 90 $90 - 140$ > 140	Low_SBP Normal_SBP High_SBP	150	$\forall hasSBP.High\_SBP$
FT	$18 - 23$ $23 - 29$ $26 - 32$	Low_FT Normal_FT High_FT	20	$\forall hasFT.Low\_FT$
SCR	$0 - 0.2$ $0.1 - 1.0$ $0.85 - 1.5$	Low_SCR Normal_SCR High_SCR	1.3	$\forall hasSCR.High\_SCR$
SCL	$0 - 2.0$ $2.0 - 22$ $20 - 25$	Low_SCL Normal_SCL High_SCL	24	$\forall hasSCL.High\_SCL$
nSRR	$0 - 2.0$ $1.0 - 3.0$ $2 - 5$	Low_nSRR Normal High_nSRR	4	$\forall has\_nSRR.High\_nSRR$
RR	$5.0 - 10$ $7.0 - 23$ $15 - 24$	Low_RR Normal_RR High_RR	12	$\forall hasRR.Normal\_RR$
Vt	$100 - 150$ $200 - 750$ $600 - 1200$	Rapid_Breath Quiet_Breath Deep_Breath	450	$\forall hasVt.Quiet\_Breath$

Table 3.2: Examples of features range and semantic description composition

Subject_1
$\forall hasANS.Normal\_Activation\_Sympathet\_Nervous\_System \sqcap \forall hasHR.Tachycardia \sqcap \forall hasSV.Low\_SV \sqcap \forall hasRos.Low\_Ros \sqcap \forall hasCO.Low\_CO \sqcap \forall hasDBP.High\_DBP \sqcap \forall hasSBP.High\_SBP \sqcap \forall hasFT.Low\_FT \sqcap \forall hasSCR.High\_SCR \sqcap \forall hasSCL.High\_SCL \sqcap \forall has\_nSRR.High\_nSRR \sqcap \forall hasEMG\_frown.EMG\_frown\_High \sqcap \forall hasEMGSmile.EMG\_Smile\_Low \sqcap \forall hasFPA.Low\_Finger\_Pulse\_Amplitude \sqcap \forall hasFPTT.Slow\_Finger\_Pulse\_Transit\_Time \sqcap \forall hasHRV.High\_HRV \sqcap \forall hasMAP.Strong\_Mean\_Arterial\_Pressure \sqcap \forall hasRR.Normal\_RR \sqcap \forall hasTe.Slow\_Expiratory\_Time \sqcap \forall hasTi.Slow\_Inspiratory\_Time \sqcap \forall hasVt.Quiet\_Breath \sqcap \forall hasTotalPower.Abnormal\_TotalPower \sqcap \forall hasPCO2.Low\_End\_Tidal\_CarbonDioxide\_Partial\_Pressure$
Emotion: Fear
$\forall hasEmotion.Fear \sqcap \forall hasANS.Normal\_Activation\_Sympathet\_Nervous\_System \sqcap \forall hasHR.Tachycardia \sqcap \forall hasSV.Low\_SV \sqcap \forall hasRos.Low\_Ros \sqcap \forall hasCO.Low\_CO \sqcap \forall hasDBP.High\_DBP \sqcap \forall hasSBP.High\_SBP \sqcap \forall hasFT.Low\_FT \sqcap \forall hasSCR.High\_SCR \sqcap \forall hasSCL.High\_SCL \sqcap \forall has\_nSRR.High\_nSRR \sqcap \forall hasEMG\_frown.EMG\_frown\_High \sqcap \forall hasEMGSmile.EMG\_Smile\_Low \sqcap \forall hasFPA.Low\_Finger\_Pulse\_Amplitude \sqcap \forall hasFPTT.Slow\_Finger\_Pulse\_Transit\_Time \sqcap \forall hasHRV.High\_HRV \sqcap \forall hasMAP.Strong\_Mean\_Arterial\_Pressure \sqcap \forall hasRR.Normal\_RR \sqcap \forall hasTe.Slow\_Expiratory\_Time \sqcap \forall hasTi.Slow\_Inspiratory\_Time \sqcap \forall hasVt.Quiet\_Breath \sqcap \forall hasTotalPower.Abnormal\_TotalPower \sqcap \forall hasPCO2.Low\_End\_Tidal\_CarbonDioxide\_Partial\_Pressure$
Hypothesys: Contentment
$\forall hasEmotion.Contentment \sqcap \forall hasRos.Normal\_Ros \sqcap \forall hasDBP.Low\_DBP \sqcap \forall hasSBP.Low\_SBP \sqcap \forall hasEMG\_smile.EMG\_smile\_High \sqcap \forall hasMAP.Weak\_Mean\_Arterial\_Pressure \sqcap \forall hasTe.Rapid\_Expiratory\_Time \sqcap \forall hasTi.Rapid\_Inspiratory\_Time \sqcap \forall hasVt.Rapid\_Breath$
Hypothesys: Anger
$\forall hasEmotion.Anger \sqcap \forall hasANS.High\_Activation\_Sympathetic\_Nervous\_System \sqcap \forall hasSV.Normal\_SV \sqcap \forall hasRos.High\_Ros \sqcap \forall hasCO.Normal\_CO \sqcap \forall hasPIPT.Rapid\_Post\_Inspiratory\_Pause\_Time$

Table 3.2 contains the complete semantic description of an example emotion individual in the Knowledge Base elicited by a trigger stimulus. The first semantic annotation intuitively shows the subject has a state of the normal activation of the sympathetic nervous system, a high blood volume pumped by the heart, a low diastolic pressure, a normal systolic blood pressure, low skin temperature, increased heart rate and medium arterial pressure during the cardiac cycle.
The output of the semantic annotation phase will be constituted by a number of annotations like this example. In detail this step automatically and dinamically builds the assertional part of KB, constituted by a number of individuals equal to the number of subjects involved in the experiment.

3.1.5 Emotion detection via semantic matchmaking

The dynamic KB creation produces the "training set" for the subsequent classification task, which exploits a semantic matchmaking process computing non-standard inferences, implemented in the Mini-ME embedded reasoning engine [116]. It was adopted to enable an advanced and automatic discovery of the most appropriate emotion for the biosignal features extracted from a test subject expressed through semantic annotations. The system is able to recognize the following emotions: Amusement, Anger, Anxiety, Contentment, Disgust, Embarrassment, Fear, Happiness, Joy, Pride, Relief, Sadness. The emotion detection problem is therefore modeled as a semantic matchmaking problem where the requests are defined by the twelve possible emotions. The resource is the semantic description constructed as the logical conjunction of concepts annotating features extracted from the unlabeled input user according to the process already explained in previous sections. The matchmaker checks the correspondence between the emotional states in the KB and user description via the following algorithm:

1. select the first Request individual from the ABox;

2. in case of compatibility between Resource and Request with respect to the reference ontology, Concept Abduction inference service is applied between the user description and the reference emotion and ranked penalty values are calculated.

3. if Resource and Request are not compatible the Concept Contraction service is applied. It detects conflicts between Resource and Request and places them in *Give up*, while *Keep* represents the compatible part or Request, on which the Concept Abduction algorithm will be applied. Penalty values for both Contraction and Abduction will be computed.

Table 3.3: Results example.

Resources	**Penalty rank (%)**
Fear	57.0
Sadness	58.0
Joy	58.0
Embarrassment	59.0
Pride	60.0
Anxiety	60.0
Amusement	61.0
Happiness	61.0
Anger	63.0
Disgust	64.0
Relief	65.0
Contentment	69.0

4. The utility function value is computed for the emotion Resource.

5. The process is repeated for each individual in the KB.

Annotations shown in the example are compatible with the test subject and the outcome of Concept Abduction produces the Hypothesis expression. Table 3.2 lists the results of matchmaking between Subject_1 and Contentment, Anger and Fear emotions. In fact, the subject has typical descriptive characteristics of an affective state of fear: high diastolic blood pressure, high heart activity and a normal activation of the sympathetic nervous system. Furthermore, fear involves temperature lowering, increased respiratory rate, a low level of CO_2 in the oxygen due to the typical effect of mouth wide open and frowning due to involuntary reaction of wide-open eyes. Concersely, Anger increases heart rate and skin temperature, while breath is rather constant and fast. The test subect also has high *EMG_frown*, while *EMG_smile* is high in the Contentment emotion.
Ranked penalty values associated with a logic-based explanation are obtained from the semantic matchmaking process. In the illustrative example, Fear is the emotion with the lowest semantic distance and it is therefore recognized by the classifier, as shown in Table 3.3.

3.2 Fuzzy learning for emotion detection

The semantic-based approach explained in the previous section constructed manually the TBox, through a literature review [66] and the aid of specific domain experts. The proposed framework evolution exploits fuzzy logic based machine learning methods to build knowledge base individuals.

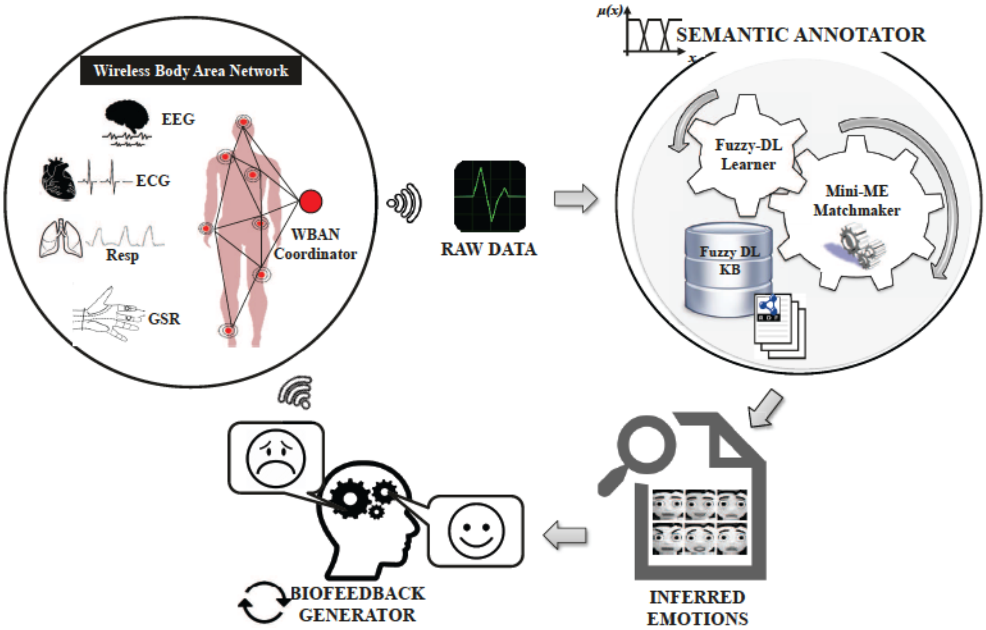


Figure 3.17: Proposed framework evolution.

The main peculiarities of the proposed approach are:

- a real-time emotional pattern detection based on *Fuzzy DLs*;
- semantic-based matchmaking to recognize the most likely emotion.

The overall architecture is depicted in Figure 3.17. Workflow phases are fundamentally the same of as the previous framework: physiological parameters are gathered and preprocessed, main features are extracted and fed into the system. The first efforts toward affect recognition have focused on finding the link between users emotional state and its corresponding physiological state, translating low-level data captured by sensing devices to high-level abstractions expressed by a semantic annotation. The goal of the new *semantic annotator* component is to build an annotation combining physiological features and bidimensional emotion parameters: *valence* and *arousal*. The key idea is to represent any emotion as a point in the valence/arousal (VA) space as proposed by Russell [109], in such a way that each affective state can be descibed as a composition of VA dimensions.

Exploiting the FuzzyDL-Learner [73, 74, 123], concept emotion descriptions are automatically learned from biosignal features. Through non-standard inference services [112], the semantic annotation is compared with emotion descriptions contained in the ontology, created in a training step from a reference biosignal dataset. Non-standard inference services for semantic matchmaking, implemented in the Mini-ME reasoner [116], produce the most appropriate elicited user emotion(s). The system can finally produce a biofeedback in order to improve emotional states.

A prototypical fuzzy ontology modeling the domain of interest was defined, using fuzzy OWL [10]. The FuzzyDL-Learner system was used to learn to identify relationship between human affective states and bidimensional emotional characteristics. The learner uses the pFOIL-DL learning algorithm [123] to automatically induce fuzzy concepts descriptions. Exploiting the FuzzyDL-Learner system [73, 74, 123] it is possible to learn automatically concept descriptions compiled into an OWL 2 [126] ontology. The main feature of FuzzyDL-Learner is learning graded fuzzy OWL 2 descriptions of a selected target class in terms of specific inclusion axioms expressed in OWL EL [88], in which fuzzy concepts may occur to improve both the accuracy and the readability of the description. The learner uses the pFOIL-DL learning algorithm [123] to automatically induce fuzzy concepts descriptions. pFOIL-DL was inspired by FOIL [101], a popular Inductive Logic Programming algorithm for learning sets of rules. The three main differences with FOIL are: (i) pFOIL-DL uses a probabilistic measure to score concept expressions, (ii) it does not remove positive examples covered from the training set, but leaves it unchanged after each learned rule and (iii) it evaluates the goodness of an induced rule not independently of previously learned rules, but considering the whole set of learnt expressions. More specifically, given a consistent knowledge base K, an OWL target class T, the *training set* ε consists of crisp concept assertions of the form $a{:}T$ where a is an individual occuring in the knowledge base and a set of fuzzy general concept inclusion (GCI) axioms $\mathcal{L}_H$, where a GCI is of the form $\langle C_1 \sqsubseteq C_2, \alpha\rangle$ (C_1 is a sub-concept of C_2 to degree at least α, $\alpha \in (0,1]$). The training set is partitioned in positive ε^+ ad negative examples ε^-. The learning problem goal is to find a finite set $H \subset \mathcal{L}_H$ of GCI, where $H = \{C_1 \sqsubseteq T, \cdots, C_n \sqsubseteq T\}$. Additionally, FuzzyDL-Learner automatically fuzzifies the range of the real-valued datatype. A discretization method, adopted by pFOIL partitions numerical datatypes into a finite number of sets (*e.g.*, *veryLow, low, medium, high, veryHigh*). Specifically, pFOIL investigates the searching the minimum and maximum value of a data property S. Then the interval $[\,min_S, max_S]$ is split into $n > 1$ equal intervals, I_1,...,I_n. Each interval has an associated fuzzy memebership function, *e.g.*, I_1 has associated a left shoulder function, I_n a right shoulder function while the middle intervals have triangular functions. The partition with $n = 5$ is usually utilized. Additionally, FuzzyDL-Learner automatically fuzzifies the range of the real-valued bidimensional emotion parameters and

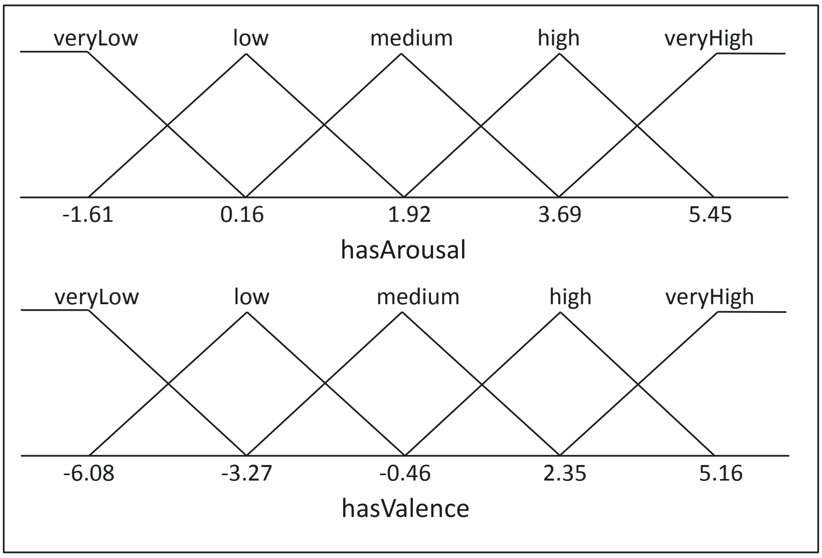


Figure 3.18: Fuzzy concepts obtained from the datatype properties *hasArousal* and *hasValence*

finds relationship between emotions and the VA space. The conjunction of the dimensional value intervals associated to each emotion –as determined by the training set– becomes the annotation for that emotion. In this way, each emotion can be described as the conjunction of qualitative features, valence and arousal. For instance, the following is a learned description for the emotion $Fear$:

$$\exists hasArousal.Arousal\_low \sqcap \exists hasValence.Valence\_high \sqsubseteq Fear$$

dictating that $Fear$ is identified by low arousal values and high valence values, where low arousal (resp. high valence) are automatically determined as illustrated in Figure 3.18. The output constitutes the *annotated dataset* and is the factual knowledge in the reference fuzzy KB.

The subsequent classification task exploits a semantic-based matchmaking process computing non-standard inferences, implemented in the Mini-ME embedded reasoning engine [116]. Like in the previous approach, semantic matchmaking was adapted to the discovery of user emotions. The request is defined by a semantic description expressed as the logical conjunction of concepts about emotional bidimensional features extracted from unlabeled user input. Resources are the semantic descriptions populating the previously annotated dataset. The matchmaker returns ranked semantic distance values associated with a logic-based explanation. The emotion with the lowest distance is identified as the best matching emotion for the current user's biosignals.

3.2.1 Fuzzy knowledge base modeling

The freely accessible DECAF [1] database for affect recognition and tagging (see Section 2.1.5) was used to assess the feasibility of the approach. The workflow starts with valence, arousal participants' self-assessment ratings information. Arousal ranges on a discrete scale of 0 (very calm) to 4 (very aroused); valence was reported on a scale from -2 (unpleasant) to 2 (very pleasant) after watching a video clip. The screenshot shown in Figure 3.20 includes a part of users' self-assessment. To make sense of the data, Z-score normalization rescaling was required. For each video normalized arousal and valence scores were calculated by taking the mean and standard deviation of arousal and valence ratings listed in [1] and shown in Figure 3.19, considered as ground truth.
In order to enable a fully automated emotion annotation and matchmaking the above meaningful emotional features must be translated using an ontology language grounded on a given logic and endowed with formal semantics. The knowledge base K namely *ABOXVA_norm_Table1.owl* was created translating normalized participants' self-assessment ratings in KB individuals. As depicted in Figure 3.21, it contains the emotional classes and a set of instances for each class, *e.g.*, *instance_216, istance_246, etc.*. For example, the annotation of *instance_288* corresponding to normalized values of the third row in Figure 3.20 is:

$$(= hasArousal 2.74) \sqcap (= hasValence 0.2) \quad \sqsubseteq \quad Fear$$

In order to tie VA parameters to emotions, the first step exploits bidimensional features as input to Fuzzy-DL Learner in order to fuzzify valence and arousal and create a Fuzzy-DL KB extracting the axioms which express each emotional label. Fuzzy-DL Learner exploits pFOIL-DL algorithm. Based on DECAF, seven emotional classes were considered to induce concept descriptions: Amusement, Anger, Disgust, Excitement, Fear, Fun and Shock. The learning problem was definded as follws: given

- an OWL target class T;
- a set of instances of the target class;

the fuzzy algorithm learns fuzzy subclass axioms of the target concept. The screenshot shown in Figure 3.22 is the FuzzyDL GUI: in the screenshot the knowledge base *ABOXVA_norm_Table1.owl* has been loaded and the first

Emotion	ID	Source Movie	L	Valence		Arousal		Scene Description
				μ	σ	μ	σ	
Amusing	01	Ace-Ventura: Pet Detective	102.1	1.22	0.53	1.03	1.00	Ace Ventura successfully hides his pets from the landlord
	02	The Gods Must be Crazy II	67.1	1.56	0.50	1.20	0.96	A couple stranded in the desert steal ostrich eggs for food
	04	Airplane	85.2	0.99	0.83	1.15	0.88	Woman and co-passengers react as pilot struggles to control aircraft
	05	When Harry Met Sally	100.2	1.05	0.61	1.08	1.02	Sally shows Harry how women fake orgasms at a restaurant
	**	Modern Times	106.4	0.87	0.69	−0.35	0.86	Bewildered factory worker in an assembly line
Funny	03	Liar Liar	55.1	0.95	0.65	0.56	0.96	Prosecution and defense discuss a divorce case in court
	06	The Gods Must be Crazy	52.1	1.26	0.56	0.81	1.15	Man tries to get past an unmanned gate on a brakeless jeep
	07	The Hangover	90.2	0.95	0.70	0.85	1.06	Group of friends on the morning after a drunken night
	09	Hot Shots	70.1	0.98	0.66	0.81	0.90	A hilarious fight sequence
Happy	08	Up	67.1	1.42	0.43	0.35	1.18	Carl—a shy, quiet boy meets the energetic Elle
	10	August Rush	90.1	0.76	0.68	−1.17	1.02	A son meets his lost mother while performing at a concert
	11	Truman Show	60.1	0.90	0.50	−1.98	0.69	Truman and his lover go to the beach for a romantic evening
	12	Wall-E	90.2	1.41	0.53	−0.82	0.91	Wall-E and Eve spend a romantic night together
	13	Love Actually	51.1	1.03	0.70	−1.38	0.80	Narrative purporting that 'Love is everywhere'
	14	Remember the Titans	52.1	0.79	0.58	−0.99	0.82	Titans win the football game
	16	Life is Beautiful	58.1	1.10	0.42	−0.16	0.79	Funny Guido arrives at a school posing as an education officer
	17	Slumdog Millionaire	80.1	0.94	0.35	−0.34	0.85	Latika and Jamal unite at the railway station
	18	House of Flying Daggers	77.2	0.84	0.56	−1.79	0.88	Young warrior meets with his love with a bouquet
Exciting	15	Legally Blonde	51.1	0.64	0.37	−0.62	0.80	Elle realizes that she has been admitted to Harvard Law School
	33	The untouchables	117.2	−0.70	0.60	1.05	0.70	Shoot-out at a railway station
Angry	19	Gandhi	108.1	−0.50	0.67	−1.00	0.92	Indian attorney gets thrown out of a first-class train compartment
	21	Lagaan	86.1	−0.98	0.49	−0.69	0.71	Indian man is helpless as a British officer threatens to shoot him
	23	My Bodyguard	68.1	−0.81	0.59	−1.35	0.79	Group of thugs provoke a teenager
	35	Crash	90.2	−1.56	0.45	0.45	0.95	A cop molests a lady in public
Disgusting	28	Exorcist	88.1	−1.52	0.64	1.71	0.90	An exorcist inquires a possessed girl
	34	Pink Flamingos	60.2	−1.95	0.61	0.18	0.83	A lady licks and eats dog faeces
Fear	30	The Shining	78.1	−0.85	0.49	1.01	0.95	Kid enters hotel room searching for his mom
	36	Black Swan	62.2	−1.07	0.35	1.00	0.73	A lady notices paranormal activity around her
	**	Psycho	76.2	−1.23	0.73	0.44	1.01	Lady gets killed by intruder in her bath tub
Sad	20	My girl	60.1	−0.85	0.62	−0.82	1.06	Young girl cries at her friend's funeral
	22	Bambi	90.1	−0.95	0.37	−0.43	1.07	Fawn Bambi's mother gets killed by a deer hunter
	24	Up	89.1	−0.99	0.45	−0.97	0.76	Old Carl loses his bedridden wife
	25	Life is Beautiful	112.1	−0.62	0.41	−0.16	0.81	Guido is caught, and shot to death by a Nazi soldier
	26	Remember the Titans	79.1	−0.84	0.53	−0.55	0.87	Key *Titans* player is paralyzed in a car accident
	27	Titanic	71.1	−0.98	0.57	−0.30	0.99	Rescuers arrive to find only frozen corpses in the sea
	31	Prestige	128.2	−1.24	0.73	1.20	0.88	Lady accidentally dies during magician's act
Shock	29	Mulholland Drive	87.1	−1.13	0.55	0.82	0.97	Man shocked by suddenly appearing frightening figure
	32	Alien	109.1	−0.99	0.71	1.22	0.76	Man is taken by an alien lurking in his room

Figure 3.19: Description of movie clips.

	Subject_ID	Video_ID	Arousal	Valence	Emotion
286	8	34	2	-2	Disgust
287	8	35	2	0	Anger
288	8	36	3	-1	Fear
289	9	1	2	-2	Amusement
290	9	2	2	0	Amusement
291	9	3	3	-2	Fun
292	9	4	1	0	Amusement
293	9	5	4	2	Amusement
294	9	6	1	1	Fun
295	9	7	3	1	Fun
296	9	9	3	-1	Fun
297	9	15	3	1	Excitement
298	9	19	1	1	Anger
299	9	21	3	1	Anger
300	9	23	2	-2	Anger
301	9	28	3	-1	Disgust
302	9	29	0	0	Shock
303	9	30	0	0	Fear
304	9	32	0	-2	Shock
305	9	33	0	0	Excitement
306	9	34	3	-2	Disgust
307	9	35	3	-2	Anger
308	9	36	3	-2	Fear
309	10	1	3	1	Amusement

Figure 3.20: The self-assessment of the users

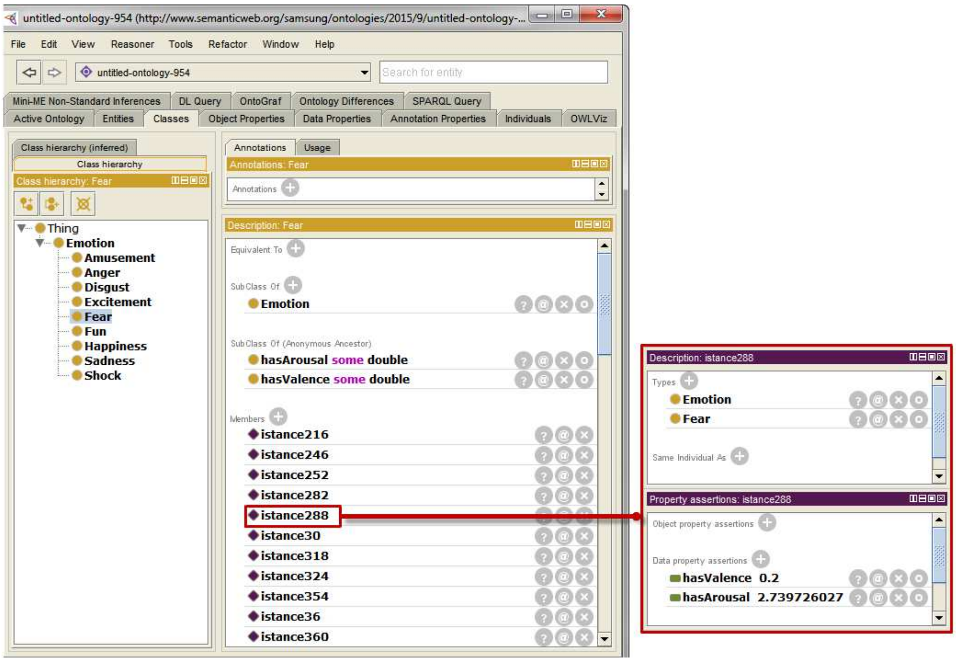


Figure 3.21: Knowledge Base

Target Class selected is *Fear*. The upper part of the screen contains information about the KB and can visualize the list of *Individuals, Concepts, Object Properties, Data Properties* and *Datatypes*. It also requires parameters like maximum depth, maximum length, θ, positive coverage, negative coverage and a flag used to assume Open World Assumption (OWA) or Closed World Assumpion (CWA). If OWA is set, it is possible to prove that an individual belongs to an atomic concept's complement. If CWA is assumed, when an individual cannot be proven to belong to a certain atomic concept then it is assumed to belong to its complement. The difference between the two assumptions lies in negative classification. The proposed approach assumes OWA, so that positive samples are all the individuals which can be proven to belong to the target class, while negative samples are the individuals which can be proven belonging to target's complement. Additionally, to guarantee termination, two parameters are set to limit the search space namely, the *maximal number of conjuncts* and the *depth of existential nestings* allowed in the left-hand side generated fuzzy GCI (setting limits implies the fact that computation may end without covering all positive samples). *Negative coverage* of a concept is defined as the percentage of negative samples covered. The goal is to keep negative coverage as low as possible. Finally, θ is a threshold parameter used to evaluate the stop algorithm criterion. In pFOIL, it is

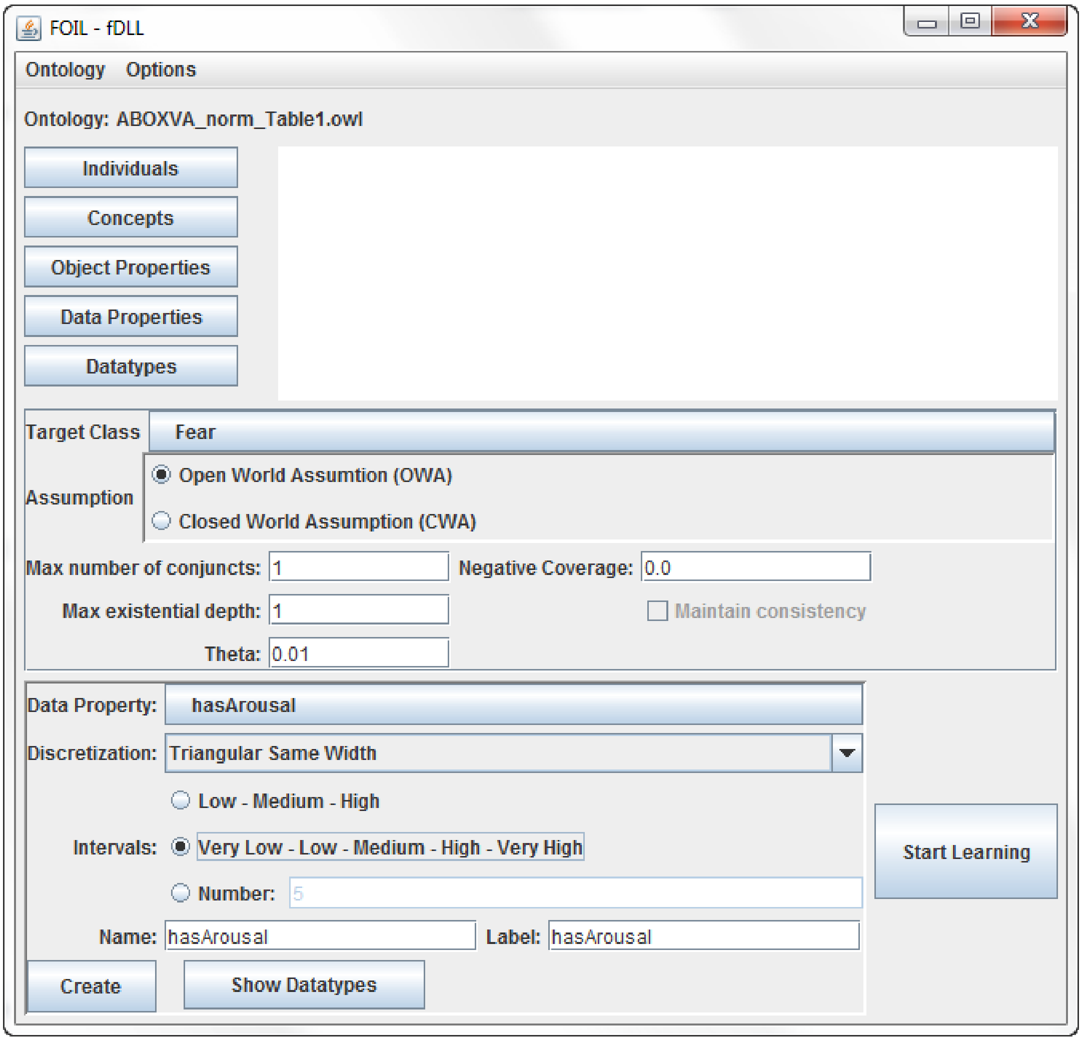


Figure 3.22: Fuzzy DL learner GUI.

imposed that no axioms are learnt unless they improve the ensemble score. So if adding a new axiom the score does not increase the axiom is not learnt. Moreover, the algorithm stops as soon as the score improvement is below the threshold θ. These parameters has been chosen manually and do not necessarily maximise performance [123]. After having described the main ingredients of pFOIL, it is possible to describe the algorithm:

1. start from concept $\top$;
2. apply a refinement operator to find more specific concept description candidates;
3. exploit a scoring function to choose the best candidate;
4. re-apply the refinement operator until a good candidate is found;
5. iterate the whole procedure until a satisfactory coverage of the positive examples is achieved.

When a data property is involved, it is possible to adopt a discretization method to partition it into a finite number of sets. After such a discretization has been obtained a fuzzy function can be defined on each set to obtain a fuzzy partition. In this work the equal width triangular partition and equal width trapezoidal partition were adopted.

They generated 5 fuzzy sets (*veryLow, low, medium, high, veryHigh*) with associated membership functions, as depicted in Figure 3.23. The datatype created for both valence and arousal were depicted in Figure 3.18. Finally, the *Start Learning* button runs the learner in order to find rules. Learnt expressions are 20 and reported in Table 3.4. This finite set of axioms composes the fuzzy KB as shown in Figure 3.24, *e.g.*, anger is characterized by high arousal and negative valence while amusement by low arousal and positive valence. The goal of this step is to build requests defined by a semantic description expressed as the logical conjunction of concepts about emotional bidimensional features extracted from unlabeled user input.

3.2.2 Annotated dataset and semantic matchmaking

As explained in the previous section, a two-step modeling was devised to tie VA parameters to emotions. The goal of the second step is to build the *annotated dataset*, connoting each valence and arousal value according to the fuzzy partitions obtained in the previous phase. For example, after viewing a video a subject reports *V=0* and *A=2*. Self-assessment valence/arousal ratings provided by participants are processed as follows:

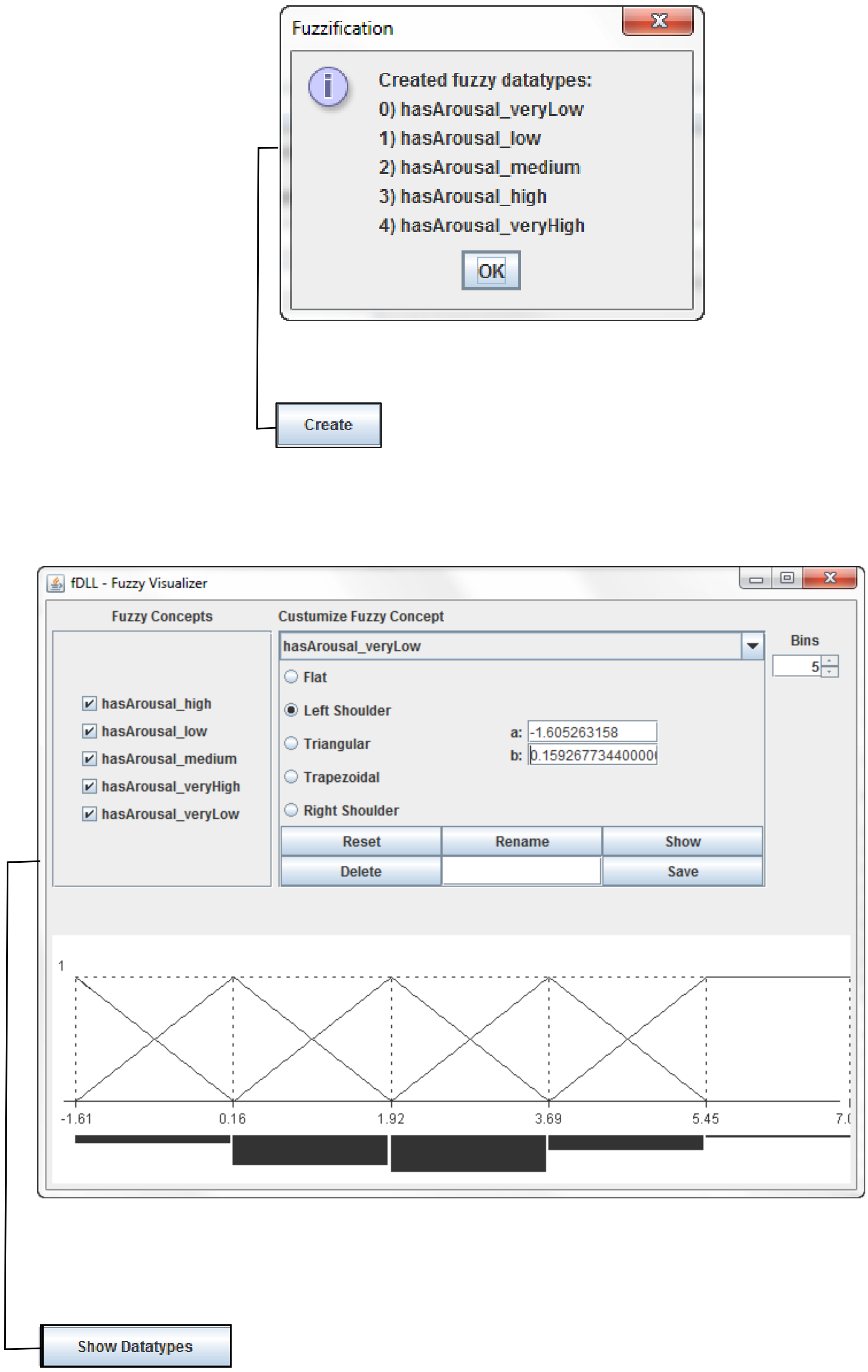


Figure 3.23: Results of discretization method.

Table 3.4: pFOIL-DL concept descriptions learned from DECAF

Target Class	**Induced Axiom**
Fear	$\exists hasArousal.Arousal\_low \sqcap \exists hasValence.Valence\_high$
	$\exists hasArousal.Arousal\_medium \sqcap \exists hasValence.Valence\_veryHigh$
	$\exists hasArousal.Arousal\_low \sqcap \exists hasValence.Valence\_veryHigh$
Amusement	$\exists hasArousal.Arousal\_low \sqcap \exists hasValence.Valence\_medium$
	$\exists hasArousal.Arousal\_low \sqcap \exists hasValence.Valence\_veryLow$
	$\exists hasArousal.Arousal\_low$
Shock	$\exists hasArousal.Arousal\_low \sqcap \exists hasValence.Valence\_high$
	$\exists hasArousal.Arousal\_veryLow \sqcap \exists hasValence.Valence\_high$
Disgust	$\exists hasArousal.Arousal\_low \sqcap \exists hasValence.Valence\_high$
	$\exists hasArousal.Arousal\_veryLow \sqcap \exists hasValence.Valence\_high$
	$\exists hasArousal.Arousal\_high \sqcap \exists hasValence.Valence\_veryHigh$
Fun	$\exists hasArousal.Arousal\_low$
	$\exists hasArousal.Arousal\_medium \sqcap \exists hasValence.Valence\_veryLow$
Anger	$\exists hasArousal.Arousal\_high$
	$\exists hasArousal.Arousal\_medium \sqcap \exists hasValence.Valence\_high$
	$\exists hasArousal.Arousal\_veryHigh$
	$\exists hasArousal.Arousal\_high \sqcap \exists hasValence.Valence\_veryHigh$
Excitement	$\exists hasArousal.Arousal\_high$
	$\exists hasArousal.Arousal\_medium \sqcap \exists hasValence.Valence\_low$
	$\exists hasArousal.Arousal\_veryLow \sqcap \exists hasValence.Valence\_high$

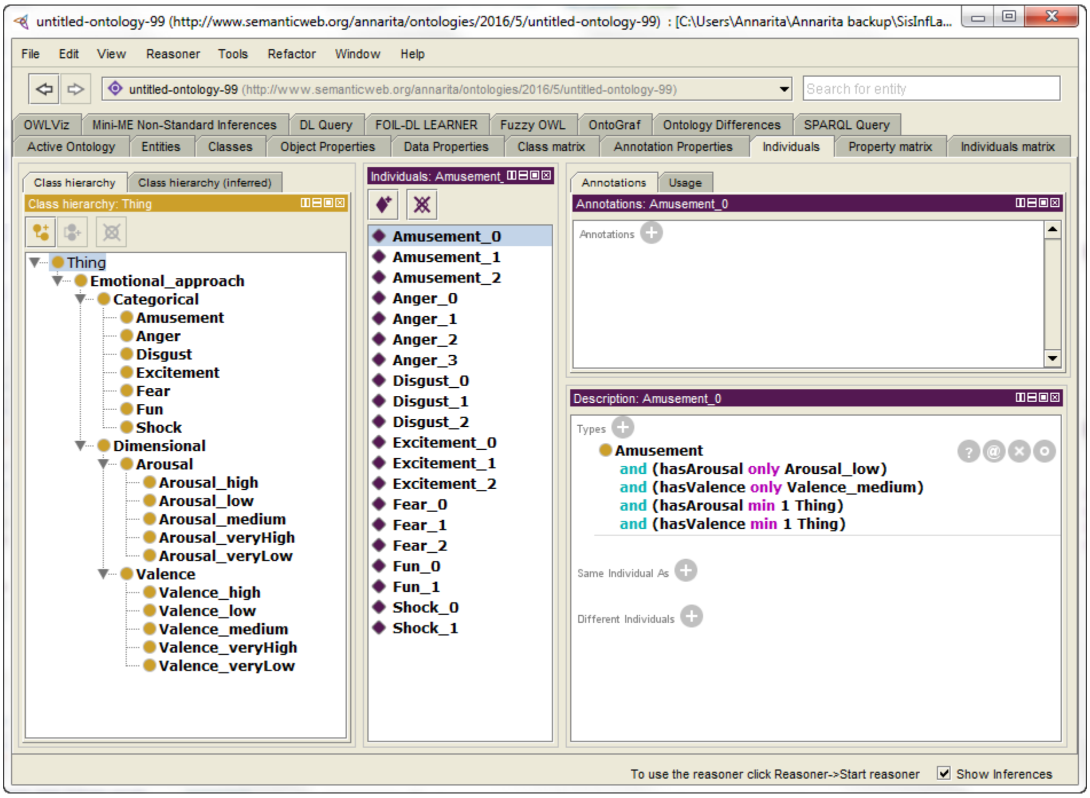


Figure 3.24: Fuzzy knowledge base.

Table 3.5: Semantic Matchmaking Results.

Resources	**Penalty**
Amusement	**16.18**
Fun	27.27
Excitement	36.36
Disgus	42.86
Fear	45.45
Shock	57.18
Anger	60.53

(1) Valence/arousal ratings are Z-score normalized considering ground truth mean and standard deviation in the training set. The chosen video clip has μ_A=1.20, σ_A=0.96, μ_V=1.56 and σ_V=0.50. The normalized ratings are $\overline{A}$*=0.83* and $\overline{V}$*=-3.12*.
(2) The semantic description of the subject according to the reference ontology is composed. According to fuzzy concepts obtained previously, normalized ratings are both in the *low* and *medium* range as shown in Figure 3.25. The system automatically choose as annotation, the GCI boby of a learned rules that fulfills the body with highest degree. Infact, applying the triangular function for $\overline{A}$*=0.83* and $\overline{V}$*=-3.12*, the membership functions are:
tri(0.83; -1.61, 0.16, 1.92)=0.619
tri(0.83; 0.16, 1.92, 3.69)=0.381
tri(-3.12; -6.08, -3.27, -0.46)=0.947
tri(-3.12; -327, -0.46, 2.35)=0.053
So, a semantic description is expressed as:
SubjectId_9: $\forall hasArousal.low \sqcap \exists hasArousal \sqcap \forall hasValence.low \sqcap \exists hasValence$

The goal of this step is to build the annotated dataset, which instances representing the resources of the system. (3) Annotated dataset and concept descriptions learned by Fuzzy-DL learner are fed to the matchmaker in order to detect the subject's emotion(s). In the case under examination, ranked penalty obtained from the semantic matchmaking process between **SubjectId_9** and the 20 semantic description learned, are shown in Table 3.5. Amusement has the lowest semantic distance and therefore the best matching emotion.

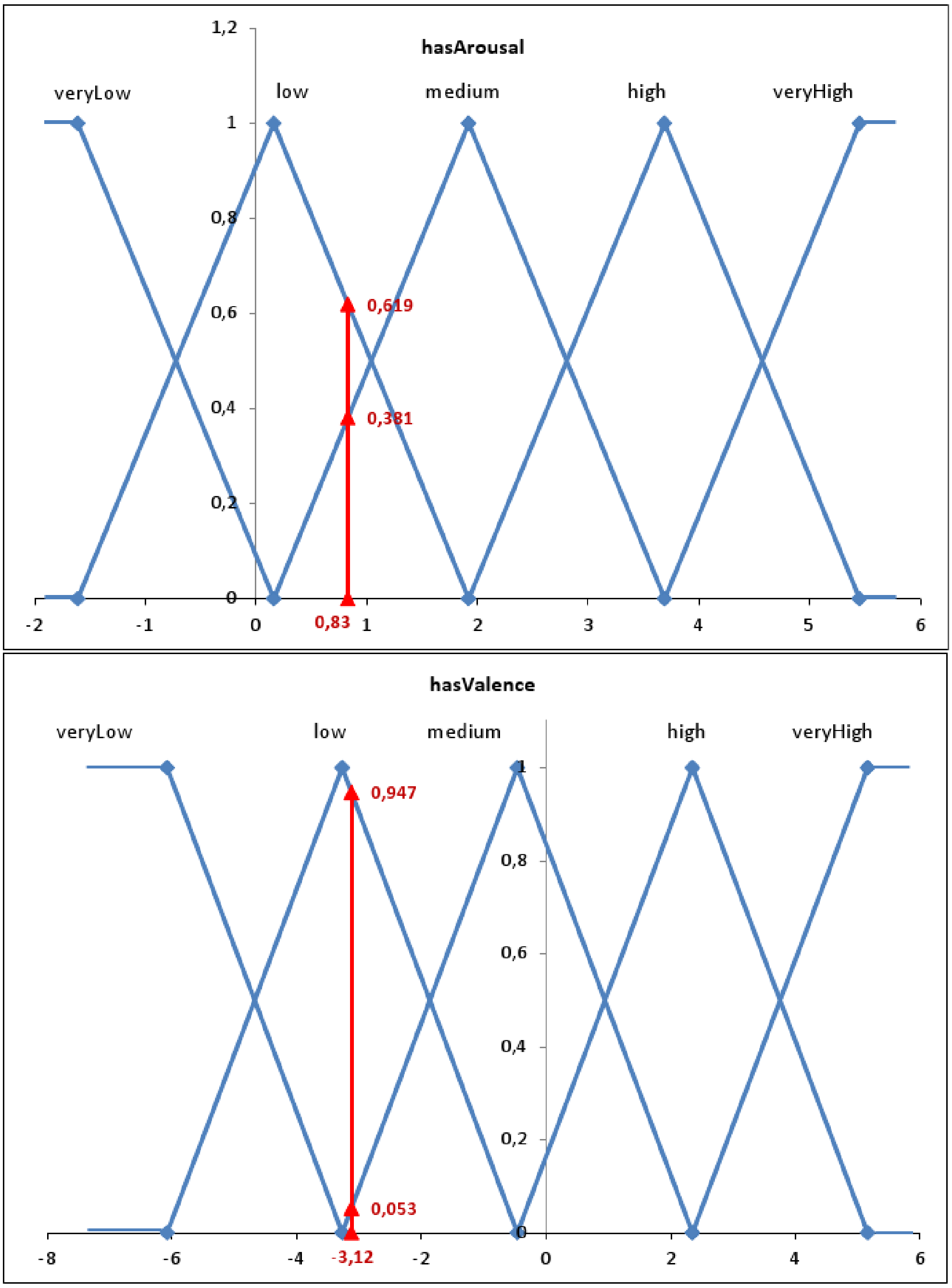


Figure 3.25: Valence/Arousal fuzzy sets

Chapter 4

Experimental campaigns

4.1 Preliminary evaluations

The freely accessible MAHNOB-HCI [119] and DEAP [65] databases for affect recognition and tagging were firstly used as ground truth. Unfortunateely, they both revealed inconsistencies in the data and a lack of information characterization. These problems prevented their use in the testing phase. In particular, a preliminary complete analysis about MANHOB-HCI has highlighted the unreliability of the dimensional data. After systematic cross-checks on the data, it was possible to assert that:

- considering one subject at a time, emotions represented according to the categorical model and the same expressed in the dimensional model often did not match. Plotting the VAD space triple on Plutchik's circumplex model, about 80% of the samples did not correspond.

- considering the whole data set, the same emotion was expressed with all allowed values in the range (from 0 to 9) for each VAD component, as shown in Table 4.1.

  This inconsistency has not allowed to use the dataset to pursue the proposed objectives. During the experimentation, its use has led to very low classification performance.

DEAP dataset contains only the dimensional space as an expression of a trigger stimulus. This was a limitation for the reliability of the results.

4.2 Syntethic benchmark experiments

This section describes the experimental setup used to assess the accuracy of the results obtained from the implementation of the knowledge-based emo-

Table 4.1: MANHOB-HCI VAD subject self-assessment

Emotion	**Valence**	**Arousal**	**Dominance**
Neutral	1-9	1-6	1-9
Anger	3-5,8	1,2,4	7-8
Disgust	1-5	3-8	1-7
Fear	1-5	3-9	1-7
Happiness	1-8	5-9	2-9
Sadness	1-8	1-4	1-9
Surprise	1-9	1-6,9	1-7,9
Amusement	1-9	2-9	1-9
Anxiety	3-9	1-6	1-8

tion recognition framework described in Section 3.1. In an effort to compensate and overcome the limitations of affective computing datasets highlighted in the previews section, a proof-of-concept evaluation of the proposed approach was carried out by generating a synthetic dataset. The main characteristics are summarized below:

- the created dataset is composed of 50 subjects, each exposed to 12 emotions, for a total of 600 instances.
- each instance contains a variable number of features; values for each feature follow a probability distribution centered around the average value in the target emotion of the instance, but with 20% probability of falling outside the typical value range.

4.2.1 Test methods

The hold out validation technique was adopted. Two thirds of the dataset instances were used as training set and the remaining third as test set. Performance measures were calculated on the test set and the model's predictive ability in order to assess the system accuracy. The 12x12 confusion matrix was plotted and accuracy, precision, recall and F-measure were computed.

4.2.2 Results and discussion

Table 4.1 shows the obtained confusion matrix. The experimental results show the proposed prototype has ranked up properly, with no false positives six emotions on twelve: Amusement, Anxiety, Embarrassment, Pride,

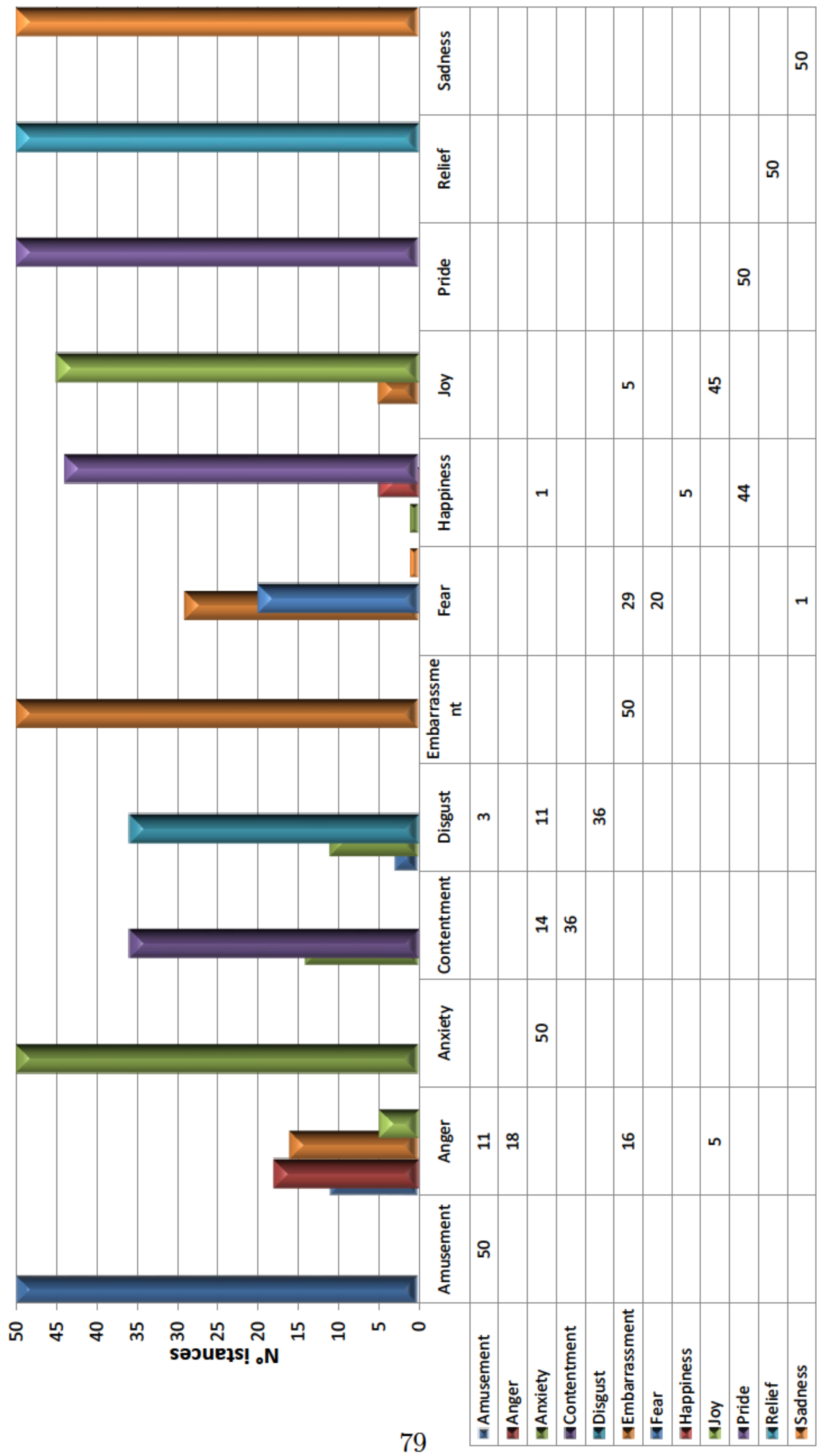
Table 4.2: Results summary on synthetic benchmark.

Emotion	**False positive**	**Correct classification (%)**
Amusement	–	100
Anger	Amusement	36
Anxiety	–	100
Contentment	Anxiety	72
Disgust	Amusement (6%), Anxiety (22%)	72
Embarrassment	–	100
Fear	Embarrassment (58%), Sadness (12%)	40
Happiness	Pride (88%), Anxiety (2%)	10
Joy	Embarrassment	90
Pride	–	100
Relief	–	100
Sadness	–	100

Relief, Sadness. The remaining six emotions have false positives. Fear and Happiness in particular were incorrectly classified in the majority of tests.

Table 4.2 provides an overview of the results in terms of the percentile rank for each emotion and the related false positives.

By analyzing false positive in greater detail, it can be said they do not represent a serious misclassification problem. Each false positive actually shares with the expected emotion at least one component of VAD space. For example, *Anger* is classified in 64% of cases as *Amusement*. This result is not senseless, because both *Anger* and *Amusement* are characterized by high arousal and dominance values. They differ for the level of valence which of course has lower value for *Anger* with respect to *Amusement*. An interesting situation occurs with *Fear*, which shares the same quadrant of the VAD plane space with its false positives: *Embarrassment* and *Sadness*. *Happiness* is classified as *Pride* in 88% of the samples, and actually they share positive valence and arousal. Overall, the system correctly predicted 76.67% of the cases, as shown in Table 4.3. Notwithstanding the inherent limitations of the test methodology, results can be deemed satisfactory and encourage further tests with properly collected datasets.



	Amusement	Anger	Anxiety	Contentment	Disgust	Embarrassment	Fear	Happiness	Joy	Pride	Relief	Sadness
Amusement	50	11			3							
Anger		18										
Anxiety			50	14	11			1				
Contentment				36								
Disgust					36							
Embarrassment		16				50	29		5			
Fear							20					
Happiness								5				
Joy		5							45			
Pride								44		50		
Relief											50	
Sadness							1					50

Figure 4.1: Confusion matrix on synthetic benchmark

Table 4.3: Classifier synthetic benchmark performance metrics.

Estimated parameters	**Value (%)**
Accuracy	76.67
Precision	86.26
Recall	91.14
F-measure	88.63

4.3 Experiments with DECAF

4.3.1 Test methods

In this case, the revised framework described in Section 3.2 was tested using the DECAF dataset, characterized by 30 subjects for 20 video clips, for a total of 600 sample instances. In this case 7 emotions, (Amusement, Anger, Disgust, Excitement, Fear, Fun ad Shock) were considered. Performance measures were the same used in the first approach. Accuracy, Recall, Precision and F-Measure have been calculated as estimated parameters.

4.3.2 Results and discussion

Table 4.4 shows the confusion matrix of emotions classification. On a total of 600 instances, 373 were correctly classified with a percentage accuracy of 62.17%. A graphical representation of the results is shown in Figure 4.2, while the overall weighted classification precion, recall and F-Measure are reported in Table 4.5.
A relevant issue is the user subjectivity associated with emotional perception: values assigned to a given impression by one person usually deviate slightly from what a different person would assign. *i.e.*, *Fun* emotion has frequently been misclassified as *Amusement* because both are pleasant and not aroused emotions. For this reason, tolerance is a crucial factor to consider in order to better convert from emotional dimensions to categories, so that each category should be represented as a region of the dimensional space, rather than just a single point. The results reveal that the semantic-based approach in conjunction with fuzzy learning is a novel research proposal with interesting perspectives.

Table 4.4: Confusion Matrix on DECAF.

Real/Predicted	**Amusement**	**Anger**	**Disgust**	**Excitement**	**Fear**	**Fun**	**Shock**
Amusement	102	6	3	4	3	0	2
Anger	5	111	0	1	3	0	0
Disgust	19	7	21	5	4	4	0
Excitement	4	15	3	36	1	0	1
Fear	1	11	2	5	40	0	1
Fun	41	21	4	5	1	46	2
Shock	11	12	2	5	8	5	17

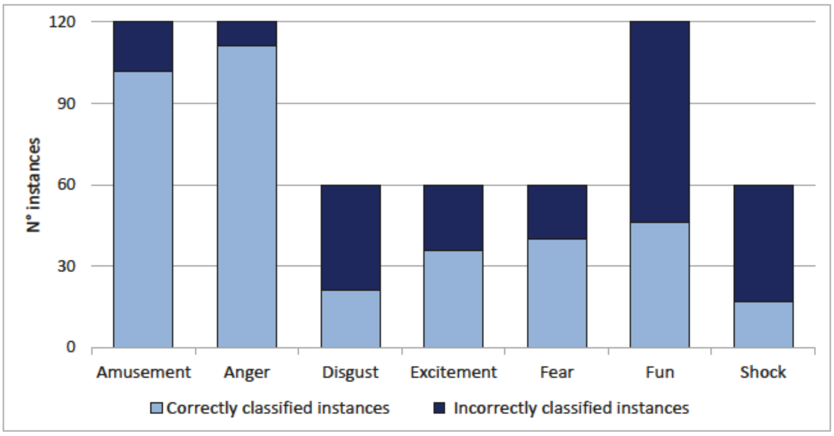


Figure 4.2: DECAF classification result.

Table 4.5: Classifier DECAF performance metrics.

Estimated parameters	**Value (%)**
Accuracy	62.17
Precision	73.5
Recall	66.0
F-measure	69.5

Chapter 5

Conclusions and perspectives

This work aimed to design, implement and test a wireless framework to classify and detect emotions. It relies on ontology languages based on Description Logics (DLs) extended to support fuzzy logic constructs. The advantage of applying semantic-based processing to raw sensor data is to make them machine-understandable and to allow for knowledge to be processed effectively, even in mobile and pervasive contexts. The proposed approach can be exploited to define innovative emotion-aware systems, which identify patterns of biosensor data and react by triggering specific actions in a given context. The perspectives of improved Affective Computing systems concern mainly potential applications in diagnosis, treatment and management of mental and stress-related disorders. The proposed approach consists of: a computing subsystem which classifies biosignals, annotates them in a logic-based formalism and performs semantic-based deductions to effectively recognize emotions. Noteworthily, the proposed computing framework only leverages off-the-shelf technology for biosignal monitoring and analysis. No specialized, expensive hardware/software facilities are needed: intrinsic complexity in emotion detection is solved by employing novel knowledge representation and learning approaches. In particular, two emotion architectures were devised. The first one is based on an Affective Computing ontology modeled from in-depth study and review of available literature, as well as the advice of domain experts. It allowed fully automatic factual knowledge generation from biosignal datasets through a rigorous mining and annotation process. The second proposal is characterized by Fuzzy DL learning in order to model emotions in the dimensional space. Both framework variants exploited semantic matchmaking in a novel way to solve the emotion recognition problem. The main scientific impact of the proposed approach consists of the adoption of new techniques and approaches for emotion recognition and representation, based on fuzzy DLs.

The main technological impacts of the research results concern the improvement of current Affective Computing applications, as well as the prototyping and industrialization of the proposed solutions. Furthermore, the increasing scale of integration of mobile and embedded infrastructures is an important direction for ICT. Increasing qualitative and quantitative performance of key software infrastructure elements such as data mining, machine learning and knowledge representation can produce benefits for a wide range of ICT applications. They can extend up to a new generation of emotionally intelligent systems for human-computer and human-robot interaction in manifold contexts. In particular, they can enable more and more intelligent Internet of Things platforms.
As future direction, the system can generate a feedback signal (physical stimulation) properly applied to avoid possible dangerous situations or even to increase the well-being of the subject. Corrective factors to be supplied as feedback stimuli must be precisely identified as they are devoted to automatically orient the system responses to improve the cognitive and emotional state of the subject. The possibility to support biofeedback in users to oppose undesirable emotions and behaviors can also have a significant impact on substance abuse and other unhealthy habits, affecting the quality of life of the general population. Finally, another interesting research perspective is referred to the analysis of the temporal dynamics of emotional states in order to build models similar to human behavior, which considers the past experiences and current needs of the individual in a given context.

Bibliography

[1] Mojtaba Khomami Abadi et al. DECAF: MEG-based multimodal database for decoding affective physiological responses. *IEEE Transactions on Affective Computing*, 6(3):209–222, 2015.

[2] Aymen A Alian and Kirk H Shelley. Photoplethysmography. *Best Practice & Research Clinical Anaesthesiology*, 28(4):395–406, 2014.

[3] Ruth Aylett, Ana Paiva, Joao Dias, Lynne Hall, and Sarah Woods. Affective agents for education against bullying. In *Affective Information Processing*, pages 75–90. Springer, 2009.

[4] Franz Baader, Diego Calvanese, Deborah L. Mc Guinness, Daniele Nardi, and Peter F. Patel-Schneider. *The Description Logic Handbook*. Cambridge University Press, 2002.

[5] John V Basmajian. Muscles alive. their functions revealed by electromyography. *Academic Medicine*, 37(8):802, 1962.

[6] Kuderna-Iulian Benta, Anca Rarău, and Marcel Cremene. Ontology based affective context representation. In *Proceedings of the 2007 Euro American conference on Telematics and information systems*, page 46. ACM, 2007.

[7] Gary G Berntson, J Thomas Bigger, Dwain L Eckberg, Paul Grossman, Peter G Kaufmann, Marek Malik, Haikady N Nagaraja, Stephen W Porges, J Philip Saul, Peter H Stone, et al. Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, (34):623–48, 1997.

[8] Franck Berthelon and Peter Sander. Emotion ontology for context awareness. In *Cognitive Infocommunications (CogInfoCom), 2013 IEEE 4th International Conference on*, pages 59–64, 2013.

[9] Katharine Blocher and Rosalind W Picard. Affective social quest. In *Socially intelligent agents*, pages 133–140. Springer, 2002.

[10] Fernando Bobillo and Umberto Straccia. Fuzzy ontology representation using OWL 2. *International Journal of Approximate Reasoning*, 52:1073–1094, 2011.

[11] Fernando Bobillo and Umberto Straccia. The Fuzzy Ontology Reasoner *fuzzyDL*. *Knowledge-Based Systems*, 95:12–34, 2016.

[12] Andrea Bonarini, Luca Mainardi, Matteo Matteucci, Simone Tognetti, and Roberto Colombo. Stress recognition in a robotic rehabilitation task. In *Robotic helpers: user interaction, interfaces and companions in assistive and therapy robotics, a workshop at ACM/IEEE HRI*, pages 41–48, 2008.

[13] Danny Oude Bos. EEG-based emotion recognition. *The Influence of Visual and Auditory Stimuli*, pages 1–17, 2006.

[14] Steven Bowbrick and Alex N Borg. *ECG complete*. Elsevier Health Sciences, 2006.

[15] John T Cacioppo, Louis G Tassinary, and Gary Berntson. *Handbook of psychophysiology*. Cambridge University Press, 2007.

[16] Rafael Calvo, Sidney D'Mello, Jonathan Gratch, Arvid Kappas, Chi-Chun Lee, Jangwon Kim, Angeliki Metallinou, Carlos Busso, Sungbok Lee, and Shrikanth S. Narayanan. Speech in affective computing. *The Oxford handbook of affective computing*, 2014.

[17] Sandra Carvalho, Jorge Leite, Santiago Galdo-Álvarez, and Óscar F Gonçalves. The emotional movie database (emdb): a self-report and psychophysiological study. *Applied psychophysiology and biofeedback*, 37(4):279–294, 2012.

[18] Ginevra Castellano, Loic Kessous, and George Caridakis. Emotion recognition through multiple modalities: face, body gesture, speech. In *Affect and emotion in human-computer interaction*, pages 92–103. Springer, 2008.

[19] Guillaume Chanel, Cyril Rebetez, Mireille Bétrancourt, and Thierry Pun. Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 41(6):1052–1063, 2011.

[20] Jing Chen, Bin Hu, Na Li, Chengsheng Mao, and Philip Moore. A Multimodal Emotion-Focused e-health Monitoring Support System. In

Seventh International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS), pages 505–510, 2013.

[21] Min Chen, Sergio Gonzalez, Athanasios Vasilakos, Huasong Cao, and Victor CM Leung. Body area networks: A survey. *Mobile networks and applications*, 16(2):171–193, 2011.

[22] Marcos E Cintra, Maria C Monard, and Heloisa A Camargo. On rule learning methods: a comparative analysis of classic and fuzzy approaches. *Soft Computing: State of the Art Theory and Novel Applications*, pages 89–104, 2013.

[23] Gari D. Clifford, Francisco Azuaje, and Patrick McSharry. *Advanced Methods And Tools for ECG Data Analysis.* Artech House, Inc., 2006.

[24] Jeffrey F Cohn and Fernando De la Torre. Automated face analysis for affective computing. *The Oxford handbook of affective computing*, page 131, 2014.

[25] Simona Colucci, Tommaso Di Noia, Agnese Pinto, Azzurra Ragone, Michele Ruta, and Eufemia Tinelli. A Non-Monotonic Approach to Semantic Matchmaking and Request Refinement in E-Marketplaces. *Int. Jour. of Electronic Commerce*, 12(2):127–154, 2007.

[26] Diane J Cook et al. Assessing the quality of activities in a smart environment. *Methods Inf Med*, 48(5):480–485, 2009.

[27] Tommaso Costa, Elena Rognoni, and Dario Galati. EEG phase synchronization during emotional response to positive and negative film stimuli. *Neuroscience letters*, 406(3):159–164, 2006.

[28] Roddy Cowie, Ellen Douglas-Cowie, Nicolas Tsapatsoulis, George Votsis, Stefanos Kollias, Winfried Fellenz, and John G Taylor. Emotion recognition in human-computer interaction. *IEEE Signal processing magazine*, 18(1):32–80, 2001.

[29] Antonio Damasio and Raymond J Dolan. The feeling of what happens. *Nature*, 401(6756), 1999.

[30] Charles Darwin. *The expression of the emotions in man and animals*, volume 526. University of Chicago press, 1965.

[31] Richard J Davidson. Affective neuroscience and psychophysiology: toward a synthesis. *Psychophysiology*, 40(5):655–665, 2003.

[32] Ellen Douglas-Cowie, Roddy Cowie, and Marc Schröder. A new emotion database: considerations, sources and scope. In *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*, 2000.

[33] Ellen Douglas-Cowie, Roddy Cowie, Ian Sneddon, Cate Cox, Orla Lowry, Margaret Mcrorie, Jean-Claude Martin, Laurence Devillers, Sarkis Abrilian, Anton Batliner, et al. The humaine database: addressing the collection and annotation of naturalistic and induced emotional data. In *International Conference on Affective Computing and Intelligent Interaction*, pages 488–500. Springer, 2007.

[34] P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. 1978.

[35] Paul Ekman. Expression and the nature of emotion. *Approaches to emotion*, 3:19–344, 1984.

[36] Paul Ekman. Basic emotions. *Handbook of Cognition and Emotion*, pages 45–60, 1999.

[37] Paul Ekman, Robert W Levenson, and Wallace V Friesen. Autonomic nervous system activity distinguishes among emotions. *Science*, 221(4616):1208–1210, 1983.

[38] Kwoting Fang and Shou-Che Lin. Exploring the relationships between colors and main meridians: An ontology-based perspective. In *2010 6th International Conference on Advanced Information Management and Service (IMS)*.

[39] Christine Fogarty and John A Stern. Eye movements and blinks: their relationship to higher cognitive processes. *International Journal of Psychophysiology*, 8(1):35–42, 1989.

[40] Virginia Francisco, Pablo Gervás, and Federico Peinado. *Ontological reasoning to configure emotional voice synthesis*. Springer, 2007.

[41] Robert R Freedman. Physiological mechanisms of temperature biofeedback. *Biofeedback and Self-regulation*, 16(2):95–115, 1991.

[42] Alan J Fridlund and John T Cacioppo. Guidelines for human electromyographic research. *Psychophysiology*, 23(5):567–589, 1986.

[43] Patrick Gebhard, Marc Schröder, Marcela Charfuelan, Christoph Endres, Michael Kipp, Sathish Pammi, Martin Rumpler, and Oytun

Türk. Ideas4games: building expressive virtual characters for computer games. In *International Workshop on Intelligent Virtual Agents*, pages 426–440. Springer, 2008.

[44] Marco Grassi. Developing HEO human emotions ontology. In *Biometric ID Management and Multimodal Communication*, pages 244–251. Springer, 2009.

[45] Michael Grimm, Kristian Kroschel, and Shrikanth Narayanan. The vera am mittag german audio-visual emotional speech database. In *2008 IEEE international conference on multimedia and expo*, pages 865–868. IEEE, 2008.

[46] Ivana Gritti, Stefano Defendi, Clara Mauri, Giuseppe Banfi, Piergiorgio Duca, Giulio Sergio Roi, et al. Heart rate variability, standard of measurement, physiological interpretation and clinical use in mountain marathon runners during sleep and after acclimatization at 3480 m. *Journal of Behavioral and Brain Science*, 3(01):26, 2013.

[47] Andreas Haag, Silke Goronzy, Peter Schaich, and Jason Williams. Emotion recognition using bio-sensors: First steps towards an automatic system. In *Affective dialogue systems*, pages 36–48. Springer, 2004.

[48] Maja Hadzic, Meifania Chen, and Tharam S Dillon. Towards the mental health ontology. In *Bioinformatics and Biomedicine, 2008. BIBM'08. IEEE International Conference on*, pages 284–288. IEEE, 2008.

[49] Janna Hastings, Werner Ceusters, Barry Smith, and Kevin Mulligan. The emotion ontology: enabling interdisciplinary research in the affective sciences. In *International and Interdisciplinary Conference on Modeling and Using Context*, pages 119–123. Springer, 2011.

[50] Jennifer A Healey and Rosalind W Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2):156–166, 2005.

[51] Jochen Heinsohn. Probabilistic description logics. In *Proceedings of the Tenth international conference on Uncertainty in artificial intelligence*, pages 311–318. Morgan Kaufmann Publishers Inc., 1994.

[52] Ikuo Homma and Yuri Masaoka. Breathing rhythms and emotions. *Experimental physiology*, 93(9):1011–1021, 2008.

[53] Matthew Horridge and Sean Bechhofer. The OWL API: a Java API for working with OWL 2 ontologies. *Proc. of OWL Experiences and Directions*, 2009, 2009.

[54] Eva Hudlicka. To feel or not to feel: The role of affect in human–computer interaction. *International journal of human-computer studies*, 59(1):1–32, 2003.

[55] William James. What is an emotion? *Mind*, (34):188–205, 1884.

[56] Eric R Kandel, James H Schwartz, Thomas M Jessell, et al. *Principles of neural science*, volume 4. McGraw-Hill New York, 2000.

[57] Ashish Kapoor, Winslow Burleson, and Rosalind W Picard. Automatic prediction of frustration. *International journal of human-computer studies*, 65(8):724–736, 2007.

[58] Preeti Khanna and M Sasikumar. Rule based system for recognizing emotions using multimodal approach. *IJACSA) International Journal of Advanced Computer Science and Applications*, 4(7), 2013.

[59] Kashif Kifayat, Paul Fergus, Simon Cooper, and Madjid Merabti. Body area networks for movement analysis in physiotherapy treatments. In *Advanced Information Networking and Applications Workshops (WAINA), 2010 IEEE 24th International Conference on*, pages 866–872. IEEE, 2010.

[60] Jonghwa Kim and Elisabeth André. Emotion recognition based on physiological changes in music listening. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(12):2067–2083, 2008.

[61] Michael Kipp and Jean-Claude Martin. Gesture and emotion: Can basic gestural form features discriminate emotions? In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–8. IEEE, 2009.

[62] Karl Christoph Klauer, Andreas Voss, and Christoph Stahl. *Cognitive methods in social psychology*. Guilford Press, 2012.

[63] Andrea Kleinsmith and Nadia Bianchi-Berthouze. Affective body expression perception and recognition: A survey. *IEEE Transactions on Affective Computing*, 4(1):15–33, 2013.

[64] R Benjamin Knapp, Jonghwa Kim, and Elisabeth André. Physiological signals and their use in augmenting emotion recognition for human–machine interaction. In *Emotion-oriented systems*, pages 133–159. Springer, 2011.

[65] Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1):18–31, 2012.

[66] Sylvia D Kreibig. Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84(3):394–421, 2010.

[67] Kyung Sup Kwak, Sana Ullah, and Niamat Ullah. An overview of IEEE 802.15.6 standard. In *2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010)*, pages 1–6. IEEE, 2010.

[68] Chin-Feng Lai, Yueh-Min Huang, Jong Hyuk Park, and Han-Chieh Chao. Adaptive body posture analysis for elderly-falling detection with multisensors. *IEEE Intelligent Systems*, 25(2):0020–30, 2010.

[69] Peter J Lang. The emotion probe: studies of motivation and attention. *American psychologist*, 50(5):372–385, 1995.

[70] Benoît Latré, Bart Braem, Ingrid Moerman, Chris Blondia, and Piet Demeester. A survey on wireless body area networks. *Wireless Networks*, 17(1):1–18, 2011.

[71] Richard S Lazarus. *Psychological stress and the coping process.* McGraw-Hill, 1966.

[72] Lei Li and Ian Horrocks. A software framework for matchmaking based on semantic web technology. *Int. Jour. of Electronic Commerce*, 8(4):39–60, 2004.

[73] Francesca A Lisi and Umberto Straccia. Learning in description logics with fuzzy concrete domains. *Fundamenta Informaticae*, 140(3-4):373–391, 2015.

[74] Francesca Alessandra Lisi and Umberto Straccia. A FOIL-Like Method for Learning under Incompleteness and Vagueness. *23rd International Conference on Inductive Logic Programming*, 8812:123–139, 2014. Revised Selected Papers.

[75] Juan Miguel López, Rosa Gil, Roberto García, Idoia Cearreta, and Nestor Garay. Towards an ontology for describing emotions. In *Emerging technologies and information systems for the knowledge society*, pages 96–104. Springer, 2008.

[76] Mohammad Malkawi and Omayya Murad. Artificial neuro fuzzy logic system for detecting human emotions. *Human-Centric Computing and Information Sciences*, 3(1), 2013.

[77] Regan L Mandryk and M Stella Atkins. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, 65(4):329–347, 2007.

[78] Jay W Mason, Douglas J Ramseth, Dennis O Chanter, Thomas E Moon, Daniel B Goodman, and Boaz Mendzelevski. Electrocardiographic reference ranges derived from 79,743 ambulatory subjects. *Journal of electrocardiology*, 40(3):228–234, 2007.

[79] Yvette Yannick Mathieu. Annotation of emotions and feelings in texts. In *International Conference on Affective Computing and Intelligent Interaction*, pages 350–357. Springer, 2005.

[80] Michael G. McKEE. Biofeedback: an overview in the context of heart-brain medicine. *Cleveland Clinic journal of medicine*, 75:S31, 2008.

[81] Gary McKeown, Michel F Valstar, Roderick Cowie, and Maja Pantic. The semaine corpus of emotionally coloured character interactions. In *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, pages 1079–1084. IEEE, 2010.

[82] Albert Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14(4):261–292, 1996.

[83] Ben Mulder, Anje Kruizinga, Arjan Stuiver, Ireen Venema, and Piet Hoogeboom. Monitoring cardiovascular state changes in a simulated ambulance dispatch task for use in adaptive automation. *Human factors in design*, pages 161–175, 2004.

[84] David G. Myers. *Theories of emotion.* Seventh Edition, New York, NY: Worth Publishers, 2004.

[85] Arturo Nakasone, Helmut Prendinger, and Mitsuru Ishizuka. Emotion recognition from electromyography and skin conductance. In *Proc. of the 5th International Workshop on Biosignal Interpretation*, pages 219–222. Citeseer, 2005.

[86] Fatma Nasoz, Kaye Alvarez, Christine L Lisetti, and Neal Finkelstein. Emotion recognition from physiological signals using wireless sensors for presence technologies. *Cognition, Technology & Work*, 6(1):4–14, 2004.

[87] Fatma Nasoz and Christine L Lisetti. Affective user modeling for adaptive intelligent user interfaces. In *International Conference on Human-Computer Interaction*, pages 421–430. Springer, 2007.

[88] OWL 2 Web Ontology Language Profiles. `http://www.w3.org/TR/2009/REC-owl2-profiles-20091027/`. W3C, 2009 (accessed Semptember 27, 2016).

[89] Maja Pantic, Michel Valstar, Ron Rademaker, and Ludo Maat. Web-based database for facial expression analysis. In *2005 IEEE international conference on multimedia and Expo*, pages 5–pp. IEEE, 2005.

[90] W Gerrod Parrott. *Emotions in social psychology: Essential readings.* Psychology Press, 2001.

[91] Amol Patwardhan and Gerald Knapp. Augmenting supervised emotion recognition with rule-based decision model. *arXiv preprint arXiv:1607.02660*, 2016.

[92] Christian Peter and Russell Beale. *Affect and emotion in human-computer interaction: From theory to applications*, volume 4868. Springer Science & Business Media, 2008.

[93] Gert Pfurtscheller and FH Lopes Da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857, 1999.

[94] Rosalind W Picard. Affective computing. *Massachusetts Institute of Technology*, 1997.

[95] Rosalind W Picard. *Affective computing.* MIT press, 2000.

[96] Rosalind W Picard and Jennifer Healey. Eight-emotion sentics data, 2002.

[97] Rosalind W Picard, Seymour Papert, Walter Bender, Bruce Blumberg, Cynthia Breazeal, David Cavallo, Tod Machover, Mitchel Resnick, Deb Roy, and Carol Strohecker. Affective learning–a manifesto. *BT technology journal*, 22(4):253–269, 2004.

[98] Rosalind W. Picard, Elias Vyzas, and Jennifer Healey. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1175–1191, 2001.

[99] Heng Yu Ping, Lili Nurliyana Abdullah, Alfian Abdul Halin, and Puteri Suhaiza Sulaiman. A study of physiological signals-based emotion recognition systems. *Int J Comput & Technol*, 11:2189–2196, 2013.

[100] Robert Plutchik and Henry Kellerman. *Theories of Emotion*. Emotion, theory, research, and experience. Elsevier Science, 2013.

[101] J. Ross Quinlan. Learning logical definitions from relations. *Machine Learning*, 5:239–266, 1990.

[102] Pierre Rainville, Antoine Bechara, Nasir Naqvi, and Antonio R Damasio. Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International journal of psychophysiology*, 61(1):5–18, 2006.

[103] Rangaraj M. Rangarayyan. *Biomedical Signal Analysis – A Case-Study Approach*. IEEE press, 2002.

[104] Pramila Rani, Changchun Liu, Nilanjan Sarkar, and Eric Vanman. An empirical study of machine learning techniques for affect recognition in human–robot interaction. *Pattern Analysis and Applications*, 9(1):58–69, 2006.

[105] Pramila Rani, Nilanjan Sarkar, and Changchun Liu. Maintaining optimal challenge in computer games through real-time physiological feedback. In *Proceedings of the 11th international conference on human computer interaction*, volume 58, 2005.

[106] Byron Reeves and Clifford Nass. *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press Cambridge, UK, 1996.

[107] Achim Rettinger, Uta Lösch, Volker Tresp, Claudia D'Amato, and Nicola Fanizzi. Mining the semantic web. *Data Min. Knowl. Discov.*, 24(3):613–662, 2012.

[108] Ma Mercedes T Rodrigo and Ryan SJd Baker. Coarse-grained detection of student frustration in an introductory programming course. In *Proceedings of the fifth international workshop on Computing education research workshop*, pages 75–80. ACM, 2009.

[109] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161–1178, 1980.

[110] James A Russell. Culture and the categorization of emotions. *Psychological bulletin*, 110(3):426, 1991.

[111] James A Russell, Maria Lewicka, and Toomas Niit. A cross-cultural study of a circumplex model of affect. *Journal of personality and social psychology*, 57(5):848–856, 1989.

[112] Michele Ruta, Eugenio Di Sciascio, and Floriano Scioscia. Concept abduction and contraction in semantic-based P2P environments. *Web Intelligence and Agent Systems*, 9(3):179–207, 2011.

[113] Juan Fernando Sánchez Rada and Carlos Angel Iglesias Fernandez. Onyx: Describing emotions on the web of data. 2013.

[114] Arman Savran, Koray Ciftci, Guillaume Chanel, Javier Mota, Luong Hong Viet, Blent Sankur, Lale Akarun, Alice Caplier, and Michele Rombaut. Emotion detection in the loop from brain signals and facial images. In *eENTERFACE 2006 Workshop*, 2006.

[115] G E Schwartz, P L Fair, P S Greenberg, J M Foran, and G L Klerman. Self generated affective imagery elicits discrete patterns of facial muscle activity. *Psychophysiology*, 12:234, 1975.

[116] Floriano Scioscia et al. A mobile matchmaker for the Ubiquitous Semantic Web. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 10(4):77–100, 2014.

[117] Ssang-Hee Seo and Jung-Tae Lee. *Stress and EEG.* INTECH Open Access Publisher, 2010.

[118] Henrique Sequeira, Pascal Hot, Laetitia Silvert, and Sylvain Delplanque. Electrical autonomic correlates of emotion. *International Journal of Psychophysiology*, 712(1):50–56, 2009.

[119] Mohammad Soleymani, Jeroen Lichtenauer, Thierry Pun, and Maja Pantic. A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, 32(1):42–55, 2012.

[120] Vince Stanford. Biosignals offer potential for direct interfaces and health monitoring. *Pervasive Computing, IEEE*, 3(1):99–103, 2004.

[121] John A Stern, Larry C Walrath, and Robert Goldstein. The endogenous eyeblink. *Psychophysiology*, 21(1):22–33, 1984.

[122] Umberto Straccia. *Foundations of Fuzzy Logic and Semantic Web Languages*. CRC Press, 2013.

[123] Umberto Straccia and Matteo Mucci. pFOIL-DL: Learning (Fuzzy) $\mathcal{EL}$ Concept Descriptions from Crisp OWL Data Using a Probabilistic Ensemble Estimation. *Proceedings of the 30th Annual ACM Symposium on Applied Computing (SAC-15)*, pages 345–352, 2015.

[124] Yun Su, Bin Hu, Lixin Xu, Hanshu Cai, Philip Moore, Xiaowei Zhang, and Jing Chen. Emotiono+: physiological signals knowledge representation and emotion reasoning model for mental health monitoring. In *Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference on*, pages 529–535. IEEE, 2014.

[125] Roy Sucholeiki and Sydney Louis. Normal EEG Waveforms, 2010.

[126] W3C OWL Working Group. OWL 2 Web Ontology Language Document Overview (Second Edition). W3C Recommendation 11 December 2012.

[127] John Yen et al. Generalizing term subsumption languages to fuzzy logic. In *IJCAI*, volume 91, pages 472–477, 1991.

[128] Lotfi A Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.

[129] Zhihong Zeng, Maja Pantic, Glenn I Roisman, and Thomas S Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE transactions on pattern analysis and machine intelligence*, 31(1):39–58, 2009.

[130] Xiaowei Zhang, Bin Hu, Jing Chen, and Philip Moore. Ontology-based context modeling for emotion recognition in an intelligent web. *World Wide Web*, 16(4):497–513, 2013.

[131] Xiaowei Zhang, Bin Hu, Philip Moore, Jing Chen, and Lin Zhou. Emotiono: an ontology with rule-based reasoning for emotion recognition. *Neural Information Processing*, pages 89–98, 2011.

List of publications

Publications in proceedings of international conferences

1. Eliana Bove, Annarita Cinquepalmi, Danilo De Filippis, Filippo Gramegna, Saverio Ieva, Giuseppe Loseto, Agnese Pinto. *A semantic-based framework for RFID-assisted port supply chains.* In **Toward Emerging Technology for Harbour sYstems and Services (TETHYS 2014 Workshop)** – July 2014.

2. Michele Ruta, Eugenio Di Sciascio, Agnese Pinto, Filippo Gramegna, Annarita Cinquepalmi. *ANDROMEDA: Adriatic IoNian and MeDiterRanean AuthOrities for Maritime SurvEillance anD Coastal and Maritime Tourism.* In **Toward Emerging Technology for Harbour sYstems and Services (TETHYS 2015)** – December 2015.

3. Michele Ruta, Floriano Scioscia, Eliana Bove, Annarita Cinquepalmi, Eugenio Di Sciascio. *A Knowledge-based Approach for Resource Discovery and Allotment in Swarm Middleware.* In **SET-222 Specialists' Meeting on 'Swarm Centric Solution for Intelligent Sensor Networks'** – June 2016.

4. Annarita Cinquepalmi, Umberto Straccia. *An Ontology-Based Affective Computing Approach for Passenger Safety Engagement on Cruise Ships.* In **The Tenth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2016)** – October 2016 (**Best Paper Award**).

5. Michele Ruta, Floriano Scioscia, Eliana Bove, Annarita Cinquepalmi, Eugenio Di Sciascio. *A Semantic-based Approach for Resource Discovery and Allocation in Distributed Middleware*, **ACM/IFIP/USENIX Middleware 2016** – December 2016.

Publications in proceedings of national conferences

1. E. Di Sciascio, D. De Venuto, S. Colucci, T. Di Noia, M. Mongiello, M. Ruta, Giuseppe Loseto, P. Nguyen, F. Scioscia, V. F. Annese, G. Basile, G. Capurso , S. Cipriani, D. De Filippis, F. Gramegna, S. Ieva, P. Pazienza, L. Rutigliani, V. Scarola, E. Bove, A. Cinquepalmi, S. Giannini, H. Ali Khattak, V. C. Ostuni, A. Pinto, J. Rosati, P. Tomeo. *SisInfLab Research Group.* In **1st WORKSHOP on the State of**

the art and Challenges Of Research Efforts at Politecnico di Bari (SCORE 2014) – December 2014.

2. Michele Ruta, Floriano Scioscia, Annarita Cinquepalmi, Silvia Cipriani, Eugenio Di Sciascio. *Dai biosegnali agli stati emotivi: un approccio semantico.* In **XVI Convegno Nazionale Associazione Italiana Ingegneri Clinici (AIIC 2016)** – April 2016

Academic activities

Summer schools

1. IOT360, *Internet of Things*. Rome, from 29-10-2014 to 01-11-2014. **First prize** at the IOT360 Hackathon event (part of IOT360 Summer School).

2. MLCI-2015, *Machine Learning*. Genoa, from 06-07-2015 to 10-07-2015.

3. RW2015, *Web Reasoning*. Berlin, from 31-07-2015 to 04-08-2015.

Seminars attended

1. *Swarm robotics research at iridia*, Prof. Marco Dorigo, Université Libre de Bruxelles, 8-04-2014.

2. *Mining user taste signals: combining recommender system*, Prof. Matthew Rowe, 10-11-2014

3. *Nuova programmazione H2020*, Dr. Alessio Gugliotta, Politecnico di Bari, 21-11-2014.

4. *Recommender Systems: an introduction*, Prof. Markus Zanker, Alpen Adria Universitat Klagenfurt, Austria, 25-11-2014.

5. *Apulian I-CiTies 2016, Laboratorio Nazionale CINI "Smart Cities and Communities"*, UniversitÃ degli Studi di Bari, 19 aprile 2015.

6. *IEEE training proposal*, Eszter Lukacs (IEEE training manager), 21-04-2015.

7. *Nuovi orizzonti per le Smart City*, Dr. Alessio Gugliotta, Politecnico di Bari, 12-06-2015.

8. *Learning fuzzy descriptions from crisp owl ontologies*, Prof. Umberto Straccia, ISTI-CNR Pisa, 14-10-2015.

9. *Programming Techniques with MATLAB*, Dr. Gareth Thomas, Academic Marketing Manager for the Education Business in MathWorks, 28-10-2015

10. *Network Physiology: from complex dynamics of individual systems to networks of organ interactions and the Human Physiolome*, Prof. Plamen Ivanov, Physics Department, Boston University and Division of Sleep Medicine Brigham and Women's Hospital and Harvard Medical School, 25-04-2016.

11. *Introduction to Robotic Operating system (ROS)*, Dr. Donato Di Paola and Dr. Antonio Petitti, CNR, ISSIA, 22/29-06-2016 e 13/27-07-2016.

12. *Technologies & Innovation: developing business creation strategies*, Workshop, Politecnico di Bari, 21-11-2016.

Speaker in international conferences

1. *Toward Emerging Technology for Harbour sYstems and Services (TETHYS 2014 Workshop)*, Bari, Italy, 15-07-2014.

2. *Toward Emerging Technology for Harbour sYstems and Services (TETHYS 2015 Workshop)*, Bari, Italy, 14-12-2015.

3. *Italian Workshop on Embedded Systems (IWES 2016 Workshop)*, Pisa, Italy, 20-09-2016.

4. *The Tenth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2016)*, Venice, Italy, 13-10-2016.

5. *ACM/IFIP/USENIX Middleware 2016*, Poster session, Trento, Italy, 14-12-2016.

Presenter in national conferences

1. *XVI Convegno Nazionale Associazione Italiana Ingegneri Clinici (AIIC 2016)*, Poster session, Bari, 7,8,9-04-2016.