1. Introduction

A new challenge in Robotics is to create systems capable of behaviour enhancement due to their interaction with humans. Research work in psychology has shown that facial expressions play an essential role in the coordination of human conversation (Boyle et al., 1994) and constitute an essential modality in human communication. Robotherapy, a field in robotics, attempts to apply the principles of social robotics to better the psychological and physiological state of the ill, the secluded, or those with physical or mental handicaps. It seems that robots can play a role of both companionship and stimulation. They must, however, be designed with a maximum of communication capacities for such a purpose. One of the first experiments in this field of robotics was carried out with elderly people in a retirement home and Paro (Shibata, 2004). These experiments clearly showed that companion robots could give a certain moral and psychological comfort to those that are most vulnerable.

In this context, the goal of the MAPH project is the realisation of a robot with the following fundamental qualities: a stuffed animal, pleasant to touch, sensors, etc. However, a robot that is too complex or too big should be avoided. The EmotiRob project, a component of the MAPH project, aims to give a robot the capacities of perception and natural language comprehension so that it can establish a formal representation of the emotional state of its interlocutor. Finally, the EmotiRob project also includes the conception of a model of the emotional states of the robot and its evolution. Following a study of the progress of research on perception and emotional synthesis, determining the most appropriate way to express emotions proved important to have a recognition rate that would be acceptable to our public. After experimentation on the subject, we have determined the minimal number of degrees of freedom necessary for a robot to express the 6 primary emotions. The second step was the definition and description of our emotional model iGrace. The experiments carried out allowed us to validate the hypotheses of the model which would be integrated into EmI - Emotional Model of Interaction. The next steps of the project will help in evaluating the robot, its expression, as well as the amount of comfort it can bring to children.

2. The MAPH project - active media for handicap

The MAPH project objective is to give comfort to vulnerable children and/or those undergoing long-term hospitalization with the help of a robot which can be used as an
emotional companion. As the use of robots in a hospital environment remains limited, we have decided to opt for simplicity in the robotic architecture, thus in the emotional expression as well. The EmotiRob project, a component of the MAPH project, aims at maintaining nonverbal interaction with children between 4 and 8 years of age. As has been shown in the synopsis (see Figure 1), the project is essentially made of three main interdependent parts:

- Recognition and understanding of a child’s spoken language.
- Emotional interaction between the child and the robot.
- Cognitive interaction between the child and the robot.

Fig. 1. General synopsis of the MAPH project

Only emotional interaction will be laid out in detail in this document.¹

3. Emotion & interaction

3.1 Definitions

3.1.1 Emotions

For several years now, human emotion has been the source of scientific research about its definition, as well as its composition. Originally, emotion is a notion of the mind, and is, therefore, analysed and studied by psychologists and physiologists. Over time, research has

¹ For more informations about others parts of the project, you can visit Website: http://www-valoria.univ-ubs.fr/emotirob/
proven that human activities are influenced by emotional states. This theory opened the door to the integration of the notion of emotion into diverse activities such as communication, negotiation, learning, etc. Moreover, computer science research naturally aims to integrate the emotional aspect in its applications for better man-machine interaction. Some of the definitions that current research is based on are given in Table 1.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descartes et al. (de Sousa, 2008)</td>
<td>Emotions are made up of primary emotions and are measured in function to a limited number of finite dimensions (e.g., level of stimulation, intensity, pleasure or aversion, one's own intention or that of others, etc.)</td>
</tr>
<tr>
<td>Magda Arnold Arnold (1960)</td>
<td>An emotion is a tendency toward an object intuitively judged as good (beneficial) or away from an object intuitively judged as bad (harmful). This attraction or aversion is accompanied by a pattern of physiological changes for approaching or avoiding.</td>
</tr>
<tr>
<td>Solomon (Solomon, 1973)</td>
<td>Emotions are judgements characterized by their temporal mode and their content evaluation; emotions are also strategic choices in the aim to protect oneself and to increase respect of oneself (inspired by Sartre’s theory, 1938 (Sartre, 1995)).</td>
</tr>
<tr>
<td>James (James, 1884) and Lange (Lange &amp; Trans)</td>
<td>Emotion is a response to physiological changes. “...we feel sorry because we cry, angry because we strike, afraid because we tremble...”.</td>
</tr>
<tr>
<td>Ortony and al. Ortony &amp; Turner (1990)</td>
<td>Emotions are valence reactions to events, agents or objects.</td>
</tr>
<tr>
<td>Greenspan (Greenspan, 1988)</td>
<td>Emotion is a conscious mental process affecting a major component of the body; it also has a lot of influence on one’s thought and action, notably to plan social interaction strategies.</td>
</tr>
<tr>
<td>Lazarus (Lazarus, 1991)</td>
<td>He highlights that appraisals are necessary and sufficient for emotion. Adding that the notion of coping allows an individual to choose strategies to confront future problems.</td>
</tr>
<tr>
<td>Scherer (Scherer, 2005)</td>
<td>In general, emotion can be seen as a sort of process that involves different parts, including subjective feeling, cognition, physical expression, the tendency of action or desires, and the neurological process.</td>
</tr>
</tbody>
</table>

Table 1. Definitions of emotions

In the design of our iGrace model, we will define emotion as a process which characterises the whole of physiological and psychological emotions of a human being for an event at a given moment. As a process, an emotion follows an algorithm which is repeated each time an event is identified:
- Perception of the event
- Reaction to the event
- Expression of the event

3.1.2 Emotional experiences

Emotions are generally characterised by subjective experiences of certain types (Parrott, 1991). Despite the complexity of the emotions and the disagreement between theoreticians on their nature, there is a consensus on the fact that subjective experience or "feeling" is an important aspect in emotion for Man. According to some theoreticians, this subjective
emotional experience or "feeling" is in part the result of internal corporal changes or the activation of "primitive" or "non-cognitive" zones of the brain. However, Parrot (Parrott, 1988) sees "feeling" as a result of activities in high level cerebral zones. Inexorably, emotional experience is also linked to these zones of activity.

Michelle Larivey (Larivey, 2002) determined four types of emotional experiences which represent our emotion type: simple emotions, mixed emotions, dismissed emotions, and pseudo-emotions. More detailed explications on the meaning and use of these different emotional experiences follow.

- Simple emotions: These are the only real emotions. They give us direct information on the state of our needs and how to satisfy them. Emotions are used to inform us of the state of our needs. Are they satisfied? To what extent? What need is it? It is important to recognize and feel our emotions. By letting the natural process of emotions happen, we are sure to be in control of the satisfaction of our needs.
- Mixed Emotions: These are defensive experiences that seem to be emotions. In fact, they are a mix of emotions and subterfuges which cause us to fool ourselves and our interlocutor. They try to "misinform" us (unlike simple emotions). Examples: guilt, jealousy, contempt, pity, shame, etc.
- Dismissed Emotions: These are usually corporally dominated experiences which happen when an emotion is dismissed or not expressed. The repressed emotion should be uncovered. Examples: worry, anxiety, feverishness, discomfort, emptiness, muscle tension, overexcitement, migraine, knot in one’s stomach, stuttering, lump in one’s throat, etc.
- Pseudo-emotions: we often mistake these for emotions, but they are concepts that give meaning to our reality, images used as metaphors, states of mind, attitudes, or assessments. In fact, they are not emotions, they are pseudo-emotions.

3.1.3 Behaviour

Like emotions, there are several different definitions of the concept of “behaviour” in the current literature (Bloch et al., 1994):

- It is a group of phenomena that can be externally observed.
- Way of being and acting of Animals and Mankind, objective manifestations of their global activity.
- Behaviour is a group of objectively observable reactions that an organism doted with a nervous system generally executes in response to stimuli from the environment, themselves being objectively observable.
- Behaviour is a reality that can be apprehended in the form of observation units and acts, the frequency and sequencing of which are likely to be modified; it translates into action the image of the situation as it is created, with its own tools, by the being that is being studied: behaviour expresses a form of representation and construction of a particular world.

Behaviour can be defined as a group of organised movements external to the organism (Castel, visité en 2009). For humans, it can be described as a group of actions and reactions (movement, physiological modification, verbal expression, etc.) of an individual in a given situation. In the rest of this paper, the behaviour of our robot will be defined as a series of actions and/or reactions to stimuli.
3.2 Founding psychological theories

3.2.1 Appraisal theory
According to the appraisal theory (Ortony et al., 1988) or the “assessment” theory, each person is always able to pick out, consciously or unconsciously, what is important to him/her in the given context (Scherer, 2005). Emotions are thus linked to one’s evaluation of the environment. Therefore, emotions are reactions to events, agents or objects. The events, agents or objects are themselves assessed in accordance to goals, norms and attitudes of the person. This theory, which is similar to a computational approach, is used in the majority of emotion modules for its generic criteria for emotion assessment. However, this process does not define the intensity of the emotions during a reaction.

3.2.2 Lazarus theory
“People are constantly evaluating relationships with the environment with respect to their implications for personal well-being.” (Lazarus, 2001)
According to Lazarus (Lazarus, 1991) there are two processes which allow the individual to stabilise his/her relationship with the environment:

• Cognitive evaluation or appraisal: adaptive process which permits conserving or modifying the relationship between the agent and his/her goals, as well as the world with its restrictions, in such a way as to maintain balance. He has determined two types of evaluation:
  - Primary evaluation: represents the pertinence of an event or the congruence of an event, or not, to the goal.
  - Secondary evaluation: represents what could or should be done in response to an event.

• Adaptation or the concept of coping: includes “cognitive and behavioural efforts to manage internal or external demands that are appraised as taxing or exceeding the resources of the person” (Lazarus & Folkman, 1984). In other words, coping is a way to adapt to difficult situations. There are two types of coping:
  - Coping centred on the problem: meaning action: we try to solve the problem or deny it to reduce its effect.
  - Coping centred on emotion: meaning thought: we do not deny the events, but make an effort to emotionally respond to the problem.

3.2.3 Scherer theory
Scherer’s component process model (Scherer, 2005) defines emotion as a sequence of changes in state in response to external or internal stimulation in relation to the interest of the individual. These changes take place in five organic systems:

• Cognitive: information processing. Evaluation of a stimulus by perception, memory, the prediction and evaluation of available information.
• Neurophysiological: change in the internal state.
• Motivational: response to the event by preparation of actions.
• Motor: expression and behaviour of the individual.
• Subjective feeling.
These components operate independently of each other during unemotional events, but work in unison in emergency situations or emotional events.
Scherer also focused on information processing and the evaluation of a stimulus. During an emotional process, the individual sequentially evaluates and event in function to a group of criteria or SECs (Sequential Evaluation Checks). These criteria are based on four main objectives which are subdivided into secondary objectives. The main criteria correspond to the most important information that the organism needs:

- **Novelty**: determine if the external or internal stimulation has changed.
- **Pleasantness**: determine if the event is pleasant or not and produce the appropriate approach or avoidance behaviour.
- **Goal significance**: determine the implications and consequences of the event. To what point will they affect my well-being or goals in the long term?
- **Coping potential**: determine if the individual is able to cope with the consequences or not.
- **Compatibility**: determine if the event is significant to personal convictions, norms, and social values.

The result of this evaluation will give the type and intensity of the emotion caused by the event. Each emotion should be able to be determined by a combination of SECs and subchecks.

### 3.2.4 Personality theory

The idea of personality remains rather complex and it is difficult to find a unanimous concept for all those who use it. The general idea bringing together the different visions is that it represents the whole of behaviours that make up an individual. Knowledge of one’s personality allows for the prediction, with a limited margin of error, of that person’s behaviour in ordinary situations, for example, professional situations. Its objective is to gain knowledge of oneself. The type of theory from analytical psychology, elaborated by psychiatrist Carl Gustav Jung (1950), defines three major characteristics of the human psyche:

- **Introversion / Extraversion**
- **Intuiting / Sensing**
- **Thinking / Feeling**

A person’s preference for one of the two poles, on the three axes, gives the psychological type. This is determined by two main personality types:

- **Introvert**
- **Extravert**

There is a second series of psychological types determined by four fundamental psychological functions that can be found in the introvert, as well as the extravert:

- **Sensing**: “S”. This process helps you obtain awareness of sensorial information and answer this information free of judgement or evaluations of it. Importance is given to experience, facts, and data.
- **Intuition**: “N”. This process, sometimes called the sixth sense, lets you perceive abstract information, such as symbols, conceptual forms and meanings.
- **Thinking**: “T”. This is an evaluation process of judgement based on objective criteria. It lets you make decisions based on rules and principles.

2 Psyche, in analytical psychology: all of the conscious manifestations of the human personality and intellect.
Feeling: "F". This process lets you make evaluations based on what is important to you, personal, interpersonal or universal values. This cognitive process of feeling evaluates situations and information subjectively.

From this, Myers and Briggs (Myers, 1987) added a dimension to Jung’s work. This dimension judges a person’s capacity for organisation and his/her aptitude in respecting the law. It added two psychological functions to those that already existed: judgement and perceiving. By reorganising these functions and preferences into four dimensions, Myers and Briggs created the Myers Briggs Type Indicator (Myers et al., 1998). The MBTI identifies 16 major personality types Cauvin & Cailloux (2005) from a pair of possible preferences for each of the four dimensions.

3.3 Computational models

3.3.1 FLAME - Fuzzy Logic Adaptive Model of Emotions
FLAME (El-Nasr et al., 2000) is a computational model of emotions based on the evaluation of events. It includes some learning components to enhance adaptation for emotion modelling. It also uses an emotion filtering component which takes motivational states into account to solve contradictory emotions. FLAME uses fuzzy logic to map events through goals to emotional intensity. The model contains three components: emotional component, learning component, and decision-making component.

3.3.2 ParLE - Adaptative Plan Based Event Appraisal Model of Emotions
ParLE (Bui et al., 2002) is a quantitative, flexible, adaptive model of emotions for a conversational agent in a multi-agent environment which has multimodal communication capacities. ParLE assesses events based on learning and a probabilistic planning algorithm. It also models personality, as well as motivational states and their role in determining the manner in which the agent experiences emotion.

Rousseau’s model of personality (Rousseau, 1996) is used in this particular model, thus classifying personality into the different processes that an agent can carry out: perceiving, reasoning, learning, deciding, acting, interacting, revealing, and feeling - all the while showing emotion. However, the model lacks specifications of the exact influence of emotions on a planning process. Furthermore, the components of models of other agents seem to make the model not quite as flexible as the authors supposed.

3.3.3 Kismet - a robot with artificial emotions
This model aims at establishing interaction between a robot, Kismet (Breazeal, 2003), and a human by using the parent-child relationship during early communication as inspiration. Cynthia Breazeal, who set out this model of emotions in 2002, placed her approach in an agent-based architecture: the different components of the system function in parallel and influence each other. This model was tested with 5 primary emotions (anger, disgust, fear, sadness, happiness) and three additional ones (surprise, interest, and excitement). The personality was not modelled because this model was inspired by the parent-child relationship.

3.3.4 Greta - The dynamics of the affective state in an animated conversational agent
Aiming to create a man-machine interface based on an animated conversation agent, C. Pelachaud and I. Poggi proposed the Greta model (de Rosis et al., 2003). Their agent model includes two closely interrelated components:
• A representation of the agent’s mind with a dynamic mechanism for updating.
• A translation of the agent’s cognitive state through facial expressions which use various available channels (gaze direction, eyebrow shape, head direction and movement, etc.). Although implementation of the personality application was created in their model, this idea was not clearly described. In other words, the real relationship between actual personality and emotion or influences of emotion on Greta’s mind is not actually identified.

3.3.5 EMA - Emotion and Adaptation
In the EMA model (Gratch & Marsella, 2005), a triggered emotion is determined from evaluation variable values such as the desirability of the event and its probability, but also by the type of agent responsible for the event and the degree of control the agent has over the situation. There is also a casual representation between events (past, present, and future) and resulting states of the agent, as well as the agent’s decision planning system which allows for the computation of variables. However, the authors do not model personality in this model.

3.3.6 GALAAD - GRAAL Affective and Logical Agent for Argumentation and Dialog
GALAAD (Adam & Evrard, 2005) is an emotional, conversational, BDI (Belief Desire Intention) agent whose architecture is based on the OCC model. Emotions influence the standard framework of a dialogue and allow for an adaptation process defined by Lazarus. The coping strategy aims at maintaining balance for the agent by reducing the intensity of negative or sensitive emotions that could cause negative effects on his/her behaviour. Nevertheless, this model has tried to integrate actual behaviour evaluation and adaptation to the architecture of the conversational agent in the dialogue game, but it does not take personality or motivational states of the emotional rational into account. Another model by Carole Adam is the PLEIAD model (Adam et al., 2007) which seems to be another version of GALAAD. In this model, the author concentrated on updating the knowledge base of the agent by introducing a logic demonstration and activation management model. Like GALAAD, PLEIAD does not integrate personality into their agent.

3.3.7 GRACE - Generic Robotic Architecture to Create Emotions
The generic GRACE model (Dang et al., 2008) defines its emotional process as a physiological emotional response triggered by an internal or external event. It is characterised by 7 components applying the appraisal, coping, Scherer, and personality theories. Being generic lets it incorporate the functionalities of all of the above-sited models. Moreover, it integrates an “Intuition” component, which does not exist in the other models, which allows it to obtain unforeseeable emotional reactions.

3.3.8 Comparison of models
Generally speaking, the three fundamental theories that characterise an emotional process are the appraisal theory, the coping theory, and the personality theory. In previous sections we have described the most useful computational models for our project. Table 2 shows model accordance with the fundamental theories.
Table 2. Comparaison of emotional models

We have chosen to instantiate and adapt the GRACE model to our project as it is the only one to apply all three theories.

4. Robotherapy

The different studies in human-robot interaction focus on two major aspects:
- Psychological robotics: studies on the behaviour between humans and robots
- Robotherapy: use of robots as therapeutic companions for people suffering from psychological or limited physiological problems.

Robotherapy is defined as a framework of human-robot interaction with the goal to reconstruct a person’s negative experiences through the development of new technological tools to create a foundation on which new positive ideas may be constructed (Libin & Libin, 2004). In other words, robotherapy offers a methodological and experimental concept which allows for the stimulation, assistance, and rehabilitation of people with physical or cognitive disorders, those with special needs, or others with physiological disabilities.

The MAPH project, which has the goal of building a robot companion, falls within the context of robotherapy for the rehabilitation and comfort of children with physical or cognitive disorders. Research has allowed for numerous robot companions having such a purpose to be created. This novel idea is based on the works carried out on a robot with a very simple architecture, but maximal expressivity.

4.1 Robots for social interaction

4.1.1 Paro

Paro (AIST, 2004) is an interactive robotic baby seal which is currently the 8th generation of a design developed by AIST in 1993. It was designed to help the elderly deal with loneliness and develop communication and affective interaction with others. It is mainly used to give companionship in Japanese retirement homes. It reacts to being touched and to the sound of voices, makes sounds, and can move its flippers, tail, as well as its eyebrows.

4.1.2 iCat

iCat (van Breemen et al., 2005) is a robot companion cat produced by the Phillips Research laboratories. It is meant to assist its user with everyday tasks such as sending messages, receiving the daily news, selecting music, pictures or videos, and even home surveillance. It can see due to cameras located behind its eyes, reacts to sound, one’s voice, as well as gestures and can express itself due to 13 servomotors.
4.1.3 Kismet
Kismet (Breazeal & Scassellati, 2000) is an expressive robot developed by MIT that has perceptual modalities. Motors are used to give it facial expressions, as well as gaze and head orientation adjustment. With 15 degrees of freedom, it is capable of expressing emotions such as surprise, happiness, anger, sadness, etc.
These motor systems let it automatically adjust its visual and auditory detectors toward the stimulus source. Four Motorola 68332 microprocessors execute the perception, motivation, behaviour, motor skills, and face motor systems. The visual system is run by nine networked 400 MHZ PCs running QNX (a real-time Unix operating system). The speech synthesis is taken on by a dual 450 MHZ PC running NT, while the speech recognition runs on a 500 MHZ PC running linux.

4.1.4 Cosmobot
Cosmobot (Lathan et al., 2005; Brisben et al., 2005) is a robot developed by AnthroTonix designed to help children with developmental and behavioural disorders. It can react to movement and voice, and can also be controlled by a child-friendly keyboard called “Mission Control”. It can repeat sentences, move parts of its body, as well as move forward and backward, and help the child during therapy. With a sensor-equipped glove a child can make the robot move or copy the machine’s movements.

5. iGrace – computational model of emotions
To realize emotional synthesis, the first step is to establish the necessary information to understand the environment (including the interlocutor). As noted above, it is important for a robot to know how to physically express itself if it is to have non-verbal communication. However, because the interlocutor is able to communicate verbally, it is necessary to understand the main words in his/her discourse, his/her intonation, as well as his/her facial expression to grasp the emotion unveiled. We would like to be able to gather the following information:
- Discourse: even though the current systems cannot perfectly understand the range of human vocabulary and especially the discourse of an interlocutor, if the context is taken into account, some words can be processed allowing for emotional reaction. The comprehension module will enable the processing of these data to be analysed, allowing us to gather a series of information, such as the subject that is doing the action, the action, the object or subject that is undergoing the action. Moreover, by combining the acts of language, times of action, and emotional state of the interlocutor, it is possible to react without completely understanding the discourse. This reaction can still be coherent with what is said.
- The sound signal: similar to the video signal, there are 2 uses for sound. The first is to reinforce the decision taken about the emotional state of the child while speaking, as prosody is not the same for the different emotions felt while speaking. The second case is for system protection. Depending on the level of sound intensity, an appropriate reaction may be taken by the system.
- The video signal: this information is useful for 3 cases. The first is to be able to follow the child’s face. During a conversation, the interlocutor could quickly be disorientated or may think that the conversation lacks interest if his/her partner is not looking at him/her. The second case allows for the affirmation of the emotional state of the child.
thanks to recognized facial expression that is associated to an emotion. Finally, the signal will be very helpful in emergency situations. As children sometimes shake their toys rather roughly, it would be necessary to stop the robot from functioning during wrongful manipulation. Thus, depending on the obstruction of the camera’s field of vision, the system can react in the appropriate way and will automatically go in standby mode if necessary to avoid any mechanical mishaps.

The second step helps to define the method to be used to react to the discourse and to have the appropriate expression. This step is crucial in that it helps to maintain interaction at its maximum if it is correctly carried out. The emotional interaction model iGrace (see Figure 2), which is based on the emotional model GRACE that we have designed, will allow us to reach our expectations. It is composed of 3 main modules (described in detail in the following subsections) which will be able to process the received information:

- "Input" Module
- "Emotional interaction" Module
- "Output" Module

Before beginning our project, we did two experimental studies. The first experiment (Le-Pévédic et al., 2006) was carried out using the Paro robot to verify if reaction/interaction with robots depended on cultural context. This experiment pointed out that there could be mechanical problems linked to weight and autonomy, as well as interaction problems linked to the robot due to lack of emotions.

The second experiment (Petit et al., 2005) was to help us reconcile the restriction of a light, autonomous robot with understanding expression capacities. Evaluators had to select the
faces that seemed to best express primary emotions among a list of 16 faces. It was one of the simplest faces that obtained the best results. With only 6 degrees of freedom (Saint-Aimé et al., 2007), it was possible to obtain a very satisfying primary emotion recognition rate.

5.1 "Input" module
This module represents the interface for communication and data exchange between the understanding module and emotional interaction module. This is where the parameters can be found, which help to obtain the information necessary to make the process go as well as possible for interaction between the child and his/her robot companion. The parameters taken into account are the following:

- Video signal: The camera, which is placed in the nose of the robot, will first help to follow the movements of the child to keep up interaction and to keep his/her attention during the interaction; this camera will also help to stop the system when the interlocutor exhibits inappropriate or unexpected behaviour.
- Sound signal: The sound system will enable understanding, as well as ensure the robot’s safety. In the case of loud screams or a panic attack assimilated to a signal that is too high, the functioning of the robot will temporarily go into standby mode. The robot will automatically generate an “isolating” behaviour.
- Actions “for the child”: They represent a group of actions that are characterised by verbs (ex. eat, sleep, play, etc.) which children most often use. These actions or verbs are put into a hierarchy or are organised in a tree structure.
- Concepts “for the child”: These are the main themes of a child’s vocabulary (ex. family, friends, school, etc.). These terms are put into a hierarchy or are organised in a tree structure with one or more levels according to the difficulty and the subtlety of reaction that we would like to create during the interaction.
- Act of language: This helps us to understand what type of discourse is being used: question, affirmation, etc., which therefore allows us to give the robot the behaviour that is most adapted to the child’s discourse. Indeed, some types of discourse, such as interrogation, require a more expressive behaviour than others.
- Tense: This lets the robot situate the discourse in time to create better interaction. The past, present, and future are implemented.
- Phase: This represents the state of mind that the child is in during the discourse. Four different phases are taken into account:
  - Imaginary: language taken from his/her imagination.
  - Real: language from real life experience.
  - Imitation: the robot should imitate the child’s behaviour.
  - Play: the robot will play a predefined game with the child.
- Emotional state: This provides information on the emotional state of the child during the discourse. It is represented by a vector of emotions giving the degree of implication or the recognition of each primary emotion (joy, fear, anger, surprise, sadness, disgust) on a scale of -1 to 2 (see Table 3).

These input values are recorded in a database which allows the robot to check if its behaviour is having a positive or negative effect on the child’s behaviour or discourse. The objective is to increase interaction time between the child and robot companion. This comparison also enables the evolution of the history, character, and personality of the robot. This step is considered as our version of coping, see “learning” or “knowledge” of the interlocutor.
Table 3. Definitions of emotion vector coefficients.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>We do not know if the emotion is implicated</td>
</tr>
<tr>
<td>0</td>
<td>Emotion not implicated</td>
</tr>
<tr>
<td>1</td>
<td>Slightly implicated emotion</td>
</tr>
<tr>
<td>2</td>
<td>Very implicated emotion</td>
</tr>
</tbody>
</table>

5.2 "Emotional interaction" module

Due to the processed input information, the robot is able to react as naturally as possible to the child’s discourse. Knowing that it is limited to only primary emotions through facial expression to maintain nonverbal discourse, we must be able to express ourselves through the other elements of the human body. To do this we decided to integrate the notion of emotional and behavioural experience into our module. The 100 emotional experiences in our database will give us a very large number of different behaviours for the model. However, we have decided, for now, to limit ourselves to only fifty entrees of emotional experience. This diversity is possible thanks to the principle of mixing emotions (Ochs et al., 2006) coupled with the dynamics of emotions (Jost, 2009). Four main elements of interaction can be found in the model:

- Moderator
- Selector of emotional experience
- Generator of emotional experience
- Behaviour

The four modules cited and described below will help us carry out the processing necessary for interaction in six steps:

1. Extraction, from list $L_1$, of emotional experiences linked to the personality of the robot – sub-module "Moderator"
2. Extraction, from list $L_2$, of emotional experiences linked to discourse – sub-module "Selector of emotional experience"
3. Extraction, from list $L_3$, of emotional experiences linked to the emotional state of the child during discourse – sub-module "Generator of emotional experience"
4. Fusion of lists $L_1$, $L_2$, and $L_3$ into $L_4$ and recalculation of coefficient associated to each emotional experience in function to:
   - Mood of the robot
   - Affect of discourse action
   - Phase and discourse act
   - Affect of the child’s emotional state
   - Affect of discourse

This process is carried out in the sub-module "Generator of emotional experience"

5. Extraction of best emotional experiences from list $L_4$ into $L_5$ – sub-module "Behaviour".
6. Expressions of emotions linked to chosen emotional experiences. These expressions determine the behaviour of the robot - sub-module "Behaviour".

5.2.1 Sub-module "Moderator"

It tells if the character, mood, personality, and history of the robot have an influence on its beliefs and behaviour. The personality of the robot, taken from the psychological definition,
is based on the MBTI model which enables it to have a list $L_1$ of emotional experiences in accordance with its personality. Currently, this list is chosen in a pseudo-random way by the robot during its initialisation. It makes a choice of 10 emotional experiences from the base which represents its profile. It is important to not select a number of emotional experiences having a negative effect higher than the number that has a positive effect. This list will be weighed in function to its mood of the day, which is the only parameter that is taken into consideration for the calculation of the coefficients $C_{emo}$ (see equation 1) of the emotional experiences. As the development is still in progress, the other parameters are not integrated into the equation used. This list will have an influence on the behaviour it is supposed to have during the discourse.

\[
C_{emo} = \begin{cases} 
\frac{1}{10} \left[ (1 - a) \times 0 + (a \times 10) \right] & \text{if positive mood} \\
\frac{1}{10} \left[ (1 - a) \times 10 + (a \times 0) \right] & \text{if negative mood} \\
0 & \text{if neutral mood}
\end{cases}
\]

with $a = \begin{cases} 
0 & \text{if negative affect} \\
1 & \text{if positive affect}
\end{cases}$

\[ (1) \]

### 5.2.2 Sub-module "Selector of emotional experience"

This module helps give the emotional state of the robot in response to the discourse of the child. The child’s discourse is represented by the list of actions and concepts that the speech understanding module can give. With this list of actions and concepts, usually represented in trio form: “concept, action, concept”, the emotional vectors $V_i$ that are associated with it can be gathered in the database. We first manually and subjectively annotated a corpus (Bassano et al., 2005) of the most common words used by children. This annotation associates an emotional vector (see Table 4) with the different words of the corpus. Each primary emotion of the vector with a coefficient $C_{emo}$ between -1 and 2 represents the individual’s emotional degree for the word. It is important to note that the association represents the robot’s beliefs for the speech and not those of the child. Actually, the annotated coefficients are statistics. However, a learning system that will make the robot’s values evolve during its lifespan is planned. The parameters that are taken into account for this evolution will mostly be based on the feedback we gather of good or bad interaction with the child during the discourse.

<table>
<thead>
<tr>
<th>Word</th>
<th>Joy</th>
<th>Anger</th>
<th>Surprise</th>
<th>Disgust</th>
<th>Sadness</th>
<th>Fear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dad (concept)</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Mum (concept)</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Sister (concept)</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Kill (action)</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Kiss (action)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Eat (action)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Coefficient of the vector (see Table 1): -1: does not know; 0: no; 1: slightly; 2: a lot**

Table 4. Extracts of emotion vectors for a list of words (action or concept)
Due to these emotional vectors, that we have combined using equation 2, it is possible for us to determine list $L_2$ of emotional experiences that are linked to the discourse. In fact, thanks to the categorisation of emotions in layers of three that Parrot (Parrott, 2000) proposes, we can associate each emotion with emotional experiences $i_{emo}$ (see Table 5). At that moment, unlike emotional vectors, emotional experiences are associated with no coefficient $C_{emo}$. However, this will be determined in function to that of the emotional vector and by applying equation 3. This weighted list, which represents the emotional state of the robot during the speech, is transmitted to the "generator".

\[
V(emo, C_{emo}) = \sum_{i=1}^{n} V_i C_{emo}
\]

with:
- $emo$: primary emotion
- $C_{emo} > 0$: emotion coefficient
- $V$: emotional vector
- $i$: word of speech index
- $n \leq 3$: number of word of speech

\[
C_{emo} = \frac{C_{emo} \times 50}{n}
\]

with:
- $C_{emo}$: emotional experience coefficient
- $C_{emo} > 0$: emotion coefficient
- $n$: number of word of speech

Table 5. Association extracts between emotions and emotional experiences

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Emotional experiences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joy</td>
<td>Happiness, contentment, enchantment, euphoria, happy, etc.</td>
</tr>
<tr>
<td>Fear</td>
<td>Panic, anxiety, fear, fright, etc.</td>
</tr>
<tr>
<td>Anger</td>
<td>Rage, agitation, detest, fury, hate, etc.</td>
</tr>
<tr>
<td>Sadness</td>
<td>Grief, deception, depression, sorrow, etc.</td>
</tr>
<tr>
<td>Surprise</td>
<td>Wonder, surprise, etc.</td>
</tr>
<tr>
<td>Disgust</td>
<td>Nausea, contempt, etc.</td>
</tr>
</tbody>
</table>

5.2.3 Sub-module "Generator of emotional experience"

This module defines the reaction that the robot should have to the child’s discourse. It is linked to all the other interaction model modules to gather a maximum amount of information and to generate the adequate behaviour(s). The information processing is done in three steps which help give a weighted emotional experience list.

The first step consists in processing the emotional state that has been observed in the child. This state is generated by a spoken discourse, prosody and will be completed in the next version of the model by facial expression recognition. It is represented by an emotional vector, similar to the one used for the words of the discourse and will have the same coefficients $C_{emo}$, which will help create a list $L_3$ of emotional experience. Coefficient $C_{emo}$ of emotional experiences is calculated by applying equation 4.

\[
C_{ emo} = C_{emo} \times 10
\]

with:
- $C_{emo}$: emotional experience coefficient
- $C_{emo} > 0$: emotion coefficient
The second step consists in combing our 3 lists (moderator($L_1$) + selector($L_2$) + emotional state($L_3$)) into $L_4$. The new coefficient will be calculated by adding it to each list for the same emotional experience (see equation 5).

$$L_{(emo, C_{emo})} = \sum_{i=1}^{n} L_i.C_{emo}$$

(5)

with

- $emo$: emotional experience
- $C_{emo} > 0$: emotional experience coefficient
- $L$: list of emotional experiences
- $i$: emotional experience index
- $n$: number of emotional experiences

The first steps carried out have first given us list $L_4$ of emotional experiences which can generate a behaviour. However, this list was created on data which corresponded to the different emotional states, as well as the discourse of the interlocutor, and the personality of the robot. Now, that have the data in hand, we will need to take into account the meaning of the discourse to find the appropriate behaviours. The goal of this third step is the recalculate the emotional experience coefficient (see Figure 3) in function to the new parameters.

![Fig. 3. Weighing of emotional experiences linked to new parameters – step 3](www.intechopen.com)
5.2.4 Sub-module ”Behaviour”
This module lets the behavioural expression that the robot will have in response to the child’s discourse be chosen. From list $L_4$, we have to extract emotional experiences with the best coefficient into a new list $L_5$. To avoid repetition, the first thing to be done was to filter the emotional experiences that had already been used for the same discourse. A historical base of behaviours associated to the discourse would help in this process. The second process is to choose $N$ emotional experiences from the list with the best coefficients. In the case of the same coefficients, a random choice will be made. We currently have set the number of emotional experiences to be extracted to three.

Another difficulty with this module is in the dynamics of behaviour and the choice of expressions. It is important not to lose the interaction with the child by constantly repeating the same expression for a type of behaviour. The choice of a large panel of expressions will help us obtain different and unexpected interaction for the same sentence or same emotional state.

5.3 ”Output” module
This module must be capable of expressing itself in function to the material characteristics it is made of: microphone/HP, motors. The behaviour comes from the emotional interaction module and will be divided into 3 main sections:

• Tone ”of voice”: characterized by a greater or lower degree of audible signal and choice of sound that will be produced by the robot. Within the framework of my research, the interaction will remain non-verbal, thus the robot companion should be capable of emitting sounds on the same tone as the seal robot ”Paro”. These short sounds based on the works of Kayla Cornale (Cornale, visited in 2007) with ”Sounds into Syllables”, are piano notes associated to primary emotions.

• Posture: characterized by the speed and type of movement carried out by each member of the robot’s body, in relation to the generated behaviour.

• Facial expression: represents the facial expressions that will be displayed on the robot’s face. At the beginning or our interaction study, we mainly work with ”emotional experiences”. These should be translated into primary emotions afterwards, and then into facial expressions. Note that emotional experience is made up of several primary emotions.

6. Operating scenario
For this scenario, the simulator and the robot will be used for expressing emotions. This system will allow us to compare the expression of the two media. The scenario takes place in 3 phases:

• System Initialization
• Simulation Event
• Processing event
• Reaction

6.1 System initialisation
At system startup, Moderator an Outputs module initialize variables like mood, personality and emotion running the robot with values in Figure 4
6.2 Simulation event

For this phase, a sentence is pronounced into the microphone allowing the system to start process. The selected phrase, extract from experiments with the robot and children in schools is: "Bouba’s mother is die". From this sentence, the team of treatment and understanding of discourse selects the following words: Mum, Be, Death. From this selection, the 9 parameters of the module Inputs will be initialized as in Figure 5.

6.3 Processing event

The emotional interaction module processes the event received and generates a reaction to the speech in six steps. Each of these steps allows us to obtain a list of emotional experiences associated with a coefficient having a value between 0 and 100.

Step 1: Personality profile

This step, performed by the sub-module Moderator, produces an initial list of responses for the robot based on its personality. The list on which treatment is based is the personality profile of the robot (see Figure 4). Applying the equation 1 at this list, we get the first list of emotional experiences $L_1$ (see Figure 4).

![Fig. 4. List $L_1$ from Moderator](image)

Step 2: Reaction to speech

This step, performed by the sub-module selector of emotional experiences, produces a list of reactions to the speech of the interlocutor. An emotional and an affect vector is associated with each concept and action of discourse, but only the emotional vector is taken into account in this step. Using the equation 2, we add the vectors coefficient for each primary common emotion. Only values greater or equal to 0 are taken into account in our calculation. In the case of joy (see Figure 5), we have: $V \cdot joie = V_1 \cdot joie + V_2 \cdot joie = 1 + 0$. This vector fusion allows us to get list $L_2$ of emotional experiences to which we apply the equation 3 to calculate the corresponding coefficients.

Step 3: Responding to the emotional state

This step, performed by the sub-module generator of emotional experiences, produces a list $L_3$ of emotional experiences for the emotional state of the speaker when the speech is done. The emotional state of the child being represented as a vector, we can obtain a list of emotional experiences to which we apply the equation 4 for coefficient.
Step 4: Fusion of lists

This step, performed by the sub-module generator of emotional experiences, allows the fusion of all lists $L_1$, $L_2$, $L_3$ into $L_4$ and computing the new coefficient of emotional experiences by using algorithm see in Figure 3. The new list $L_4$ is see in Figure 7.

Step 5: Selection of the highest coefficients

This step, performed by the sub-module behavior, achieves the 3 best emotional experiences of the list $L_4$ into $L_5$. The list will be first reduced by deleting emotional experiences that have already been chosen for the same speech. In the case of identical coefficients, a random selection will be made.

www.intechopen.com
Fig. 7. From Generator to Output module – List $L_4$ and $L_5$

**Step 6: Initialization parameters of expression**

The last step, performed by the sub-module behavior, calculates the parameters for the expression of the reaction of the robot. We obtain the time expression in second of each emotional experience (see Figure 7).

6.4 Reaction

This last phase, carried out by the output module, simulates the robot’s reaction to the speech. With the list $L_5$ (see Figure 7) of reaction given by the emotional interaction module. For each of the emotional experiences of the list associated with one or more emotions, we randomly choose a facial expression in the basic pattern. This will be expressed using the motor in the case of the robot or the GUI in the case of the simulator.

7. Experiments

The goal of the first experiment was to partially evaluate and validate the emotional model. For this, we start experiment with a small public of all ages to gather the maximum amount of information on the improvements needed for interaction. After analysis of the results, the first improvements were made. For this experiment, only the simulation interface was used.

7.0.1 Protocol

For the first step, having been carried out among a large public, it was not difficult to find volunteers. However, we limited the number to 10 people because as we have already stated, this is not the targeted public. We did not want to modify the interaction in function to remarks made by adults. The first thing that was asked was to use abstraction as the interface represented the face and behaviour of the robot, and that the rest (type of input, ergonomy, etc.) was not to be evaluated. Furthermore, these people were asked to put themselves in the place of a targeted interlocutor so as to make the most useful remarks.

To carry out the tests, we first chose a list of 4 phrases upon which the testers were to base themselves. For each one, we included the following language information:

- Time of action: present.
- Language act: affirmative.
- Discourse context: real life.

This system helped us gain precious time that each person would use to make their decisions. The phrases given were the following:

- Mum, Hug, Dad.
• Tiger, Attack, Grandma.
• Baby, Cry.
• I, Tickle, Sister.

7.1 Evaluation grid
After the distribution and explication of the evaluation grids, each person first had to go through the following steps:
1. Give an affect (positive, negative, or neutral) to each word of the phrase.
2. Define their emotional state for the discourse.
3. Predict the emotional state of the robot.

Although this step was easy to do, it was rather long to input because some people had trouble expressing their feelings. After inputting the information we could start the simulation for each phrase. We asked the users to be attentive to the robot’s expression because it could not be seen again. After observation of the robot’s behaviour, the users had to complete the following information:
1. Which feelings could be recognized in the behaviour, and what was their intensity on the scale: not at all, a little, a lot, do not know.
2. The average speed of the expression and length of the behaviour on a scale: too slow, slow, normal, fast, too fast.
3. Did you have the impression there was a combination of emotions? Yes or no?
4. Was the sequence of emotions natural? Yes or no?
5. Are you satisfied with the robot’s behaviour? Not at all, a little, very much?

7.2 Results
The objective of this experiment was to evaluate the recognition of emotions through the simulator, and especially to determine if the response the robot will give to the speech was satisfying or not. As regards the rate of appreciation of the behaviour for each speech, 54% for a lot of satisfaction and 46% for a little, we observed that all the users found the simulator’s response coherent, and thereafter admitted that they would be fully satisfied if the robot was as they were expected. The fact that testers answered about the expected emotions had an influence on overall satisfaction.

For the rate of emotions recognition, 82% in average, the figures were very satisfactory and allowed us to prepare the next evaluation on the classification of facial expressions for each primary emotion. Not all emotions are on the graph because they bore no relation to the sentences chosen. We have also been able to see that even if the results were still rather high, there were some emotions which were recognized although they were not expressed. This confirms the need to classify, and especially the fact that each expression can be a combination of emotions. The next question is to know if the satisfaction rate will be the same with the robot after the integration of the emotional model. The other results were useful for the integration of the model on the robot:
• Speed of expressions: normal with 63%
• Behaviour length: normal with 63%
• Emotional combination: yes with 67%
• Natural sequences: yes with 71%
8. EmI - robotic conception

EmI is currently in the integration and test phase for future experiments. This robot was partially conceived by the CRIIF for the elaboration of the skeleton and the first version of the covering (see Figure 8(c)). The second version (see Figure 8(d)), was made in our laboratory. We will briefly present the robotic aspect of the elaborated work while waiting for the second generation of it.

![Head conception](image1)
![Body conception](image2)

![EmI version 1](image3)
![EmI version 2](image4)

Fig. 8. EmI conception

The skeleton of the head (see Figure 8(a)) is completely made of ABS and contains:

- 1 camera at nose level to follow the face and potentially for facial recognition. The camera used is a CMUCam 3.
- 6 motors creating the facial expression with 6 degrees of freedom. Two for the eyebrows, and four for the mouth. The motors used are AX-12+. This allows us to communicate digitally, and soon with wireless thanks to Zigbee, between the robot and a distant PC. Communication with the PC is done through a USB2Dynamixel adapter using a FTDI library.

The skeleton (see Figure 8(b)) of the torso is made of aluminium and allows the robot to turn its head from left to right, as well as up and down. It also permits the same movements at the waist. There are a total of 4 motors that create these movements.
Currently, communication with the robot is done through a distant PC directly hooked up to the motors. In the short term, the PC will be placed on EmI to process while allowing for interaction. The PC used will be a Fit PC Slim, at 500 Mhz, with 512 Mo of RAM and a 60 Go hard drive. The exploitation system used is Windows XP. It is possible to hook up a mouse, keyboard, and screen for modifications and to make the system evolve at any moment.

9. Conclusion and perspectives

The emotional model iGrace we propose allows to react emotionally to a speech given. The first experiment conducted on a small scale has enabled us to answer some questions such as length and speed of the robot expression, methods of information processing, consistency of response and emotion recognition on a simulator. To fully validate the model, a new large-scale experimentation will be repeated.

The 6 degrees of freedom used for the simulation give recognition rate very satisfactory. It is our responsibility now to make a similar experiment on the robot to evaluate its expressiveness. In addition, we undertook extensive research on the dynamics of emotions in order to increase the fluidity of movement and make the interaction more natural. The second experiment, with the robot, will allow to compare the recognition rate between the robot and the simulator.

The next version of EmI will integrated a new texture, camera recognition and prosody treatment. These parameters will allows us have a best recognition for emotional state of the child. Some parts of modules and su-modules of the model have to be develop for a best interaction.

10. Acknowledgements

EmotiRob is a project that is supported by ANR through the Psirob programme. The MAPH project is supported by regional funding from la rgion Martinique and la rgion Bretagne. We would like to first of all thank the different organisations for their financial support as well as their collaboration.

The authors would also like to thank all of the people who have contributed to the evaluation grids for the experiments, as well as the members of the Kerpape centre and IEA "Le Bondon" centre for their cooperation.

Finally, the authors would also like to thank all of the participants in the experiments for their time and constructive remarks.

11. References


AIST (2004). Seal-type robot “paro” to be marketed with best healing effect in the world.

URL: http://www.aist.go.jp/aist/e/latest research/2004/20041208 2/20041208 2.html


URL:http://www.google.com/hostednews/afp/article/ALeqM5jBCTtjOmxw1ZAGOJaWNXK6itOmsA

URL: http://psychobiologie.ouvaton.org/


URL: www.soundsintosyllables.com


Lange, C. G. & (Trans), I. A. H. (1922). *The emotions*, Williams & Wilkins Co, Baltimore, MD, US.


www.intechopen.com
Rapid advances in the field of robotics have made it possible to use robots not just in industrial automation but also in entertainment, rehabilitation, and home service. Since robots will likely affect many aspects of human existence, fundamental questions of human-robot interaction must be formulated and, if at all possible, resolved. Some of these questions are addressed in this collection of papers by leading HRI researchers.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: