

humaine

D7c

**Description of potential exemplars:
Emotion in Cognition and Action**

Lola Cañamero and WP 7 members



Version 1.0

Date: 30th November 2004

IST project contract no.	507422
Project title	HUMAINE Human-Machine Interaction Network on Emotions
Contractual date of delivery	<i>Month 11</i>
Actual date of delivery	<i>November 30, 2004</i>
Deliverable number	D7c
Deliverable title	Description of potential exemplars: Emotion in Cognition and Action
Type	Report
Number of pages	38
WP contributing to the deliverable	WP 7
Task responsible	Lola Cañamero, UH
Author(s)	Lola Cañamero, Fiorella de Rosis, Paolo Petta, Sabine Payr and other WP7 members (see pages 5-6)
EC Project Officer	Philippe Gelin

Address of lead author:

Lola Cañamero
*School of Computer Science
 University of Hertfordshire
 College Lane
 Hatfield, Herts AL10 9AB
 United Kingdom*

Table of Contents

THE PLACE OF THIS REPORT WITHIN HUMAINE	5
1 BRIEF OVERVIEW OF THE WORKPACKAGE, THE EXEMPLAR PROPOSAL, AND RELEVANT RESOURCES	7
2 RATIONALE FOR THE EXEMPLAR PROPOSAL	12
3 TECHNICAL ASPECTS OF THE PROGRAM OF RESEARCH	16
4 PROPOSED ALLOCATION OF TASKS.....	24
5 PROVISIONAL TIMELINE	38
6 CONCLUSION AND WAY FORWARD.....	42

The place of this report within HUMAINE

The HUMAINE Technical Annex identifies a common pattern that is followed by most of the project's workpackages:

“The measure of success will be the ability to generate a piece of work in each of the areas which exemplifies how a key problem in the area can be solved in a principled way, and which also demonstrates how work focused on that area can integrate with work focused on the other areas. We call these pieces of work *exemplars*. The exact form of an exemplar is not pre-specified: it may be a working system, but it might also be a well-developed design, or a representational system, or a method for user-centered design.” (p 4)

“To that end, each thematic group will work out a proposal for common action, embodied in one or more exemplars to be built during the second half of the funding period.” (p.16)

“The process will begin with production by each thematic group of a review of key concepts achievements and problems in its thematic area; and drawn from the review, an assessment of the key development goals in the area. This review and assessment will be circulated to the whole network for discussion and comment, aimed both at building understanding of basic issues across areas, and at identifying the choices of goal that would be most likely let the different groups achieve complementary developments. That consultation phase will provide the basis for deliverables in month 11, which describe in some detail a few alternatives that might realistically be chosen as exemplars in each area, and their linkages to issues in other thematic areas. A decision and planning period will follow, involving consultation within and between thematic areas, leading to presentations at the second plenary conference, which will describe a single exemplar that has been chosen for development in each area, and the way work on the exemplar will be divided across institutions. The remainder of the project will be absorbed in developing the chosen exemplar.” (p. 21)

The review and assessment documents were delivered in May, and the consultation phase has been ongoing since, using several channels, notably e-mail exchanges moderated by the coordinator; meetings of workpackage leaders by teleconferencing; meetings between workpackage representatives attending the WP4 workshop, the Summer School, and other non-HUMAINE events, such as the recent workshop “Dimensions of Sociality: Shaping Relationships with Machines” hosted by OFAI on November 19 & 20, 2004; and a consultation meeting of WP leaders and other WP representatives in Paris on October 29th & 30th, 2004.

This deliverable is one of the group of deliverables arising from the consultation phase. The function of this phase is defined as to “describe in some detail a few alternatives that might realistically be chosen as exemplars in each area”. In general, we believe that we have progressed more quickly than we have expected, and that the alternatives described here are close to the ones that should be pursued. What remains to be completed is largely detailed planning. Given the intricacy of the network, that is not a trivial task.

The following **institutions** and *persons* have contributed to the deliverable:

UH: *Lola Cañamero, Orlando Avila-García, Arnaud Blanchard, Rene te Boekhorst, Sue Attwood*

OFAI: *Paolo Petta, Sabine Payr, Stefan Rank, Bernhard Jung*

DI-BARI: *Fiorella de Rosis, Valeria Carofiglio*

UNIGE-GERG: *Klaus Scherer, David Sander, Etienne Roesch*

UNIGE-MIRALab: *Arjan Egges*

EPFL: *Etienne de Sevin, Daniel Thalmann*

Paris8: *Chrystopher Peters, Catherine Pelachaud*

HW: *Ruth Aylett*

INESC-ID: *João Dias, Ana Paiva*

IST: *Carlos Martinho*

KCL: *John Taylor, Peter Goldie, Sabine Dörner*

DIST: *Gualterio Volpe*

DFKI : *Patrick Gebhard, Marc Schröder*

ISSC-NTUA: *Kostas Karpouzis*

France Telecom RD: *Franck Panaget*

QUB: *Roddy Cowie, Edelle McMahon*

UA: *Elisabeth André*

CNRS STIC-SDV EPML 38 (new member): *Jacqueline Nadel*

ISTC (new member): *Cristiano Castelfranchi, Isabella Poggi, Maria Miceli*

USC (new member): *Stacy Marsella*

1 Brief overview of the workpackage, the exemplar proposal, and relevant resources

1.1 The field covered by the workpackage

The area covered by this workpackage is described in the Technical Annex, particularly in Section 6.2, and in more depth in the review and assessment document for the workpackage, D7b. We summarize the area here partly to make the deliverable readable as a stand-alone document, and partly to draw attention to changes of emphasis that have taken place during the first period of HUMAINE.

1.1.1 Conception of the area before HUMAINE began

Our conception of research on computational models of emotions in cognition and action before HUMAINE began was reflected in the Technical Annex, and we summarize it briefly here.

Nowadays the different disciplines concerned with the study of emotion agree on the view that, in humans, emotions entail distinctive integrated ways of perceiving and assessing situations, processing information, and modulating and prioritizing actions. Recent research in psychology and neuroscience has provided evidence that emotions pervade human intelligence at many levels, being inseparable from cognition and action. Perception, attention, memory, learning, motivation, decision-making, task (behavior) execution, adaptation, regulation of survival-related internal bodily functions (e.g., homeostasis) and the regulation of our interactions with the (physical and social) external environment are some of the aspects influenced by emotions, in addition to the features of expression, social interaction and communication investigated under other workpackages of HUMAINE. Such findings have led to a paradigm shift, departing from the rationalistic tradition that equated intelligence with pure reasoning and that considered emotions as undesirable consequences of our embodiment that hampered reasoning. On the contrary, emotions are nowadays regarded as a necessary component of cognition and intelligent behavior, and therefore they offer a rich potential for the design of artificial intelligent and interactive systems and for enhancing our interactions with them.

Elaborating computational models that embed the effects of emotions in cognition and action is a complex, multi-faceted problem that presents considerable integration challenges at various levels. Theories and models of human emotions are diverse and normally address specific aspects of the involvement of emotion in cognition and action. The same problem re-occurs in existing computational models, which in addition need a more sound understanding of their theoretical foundations and of how theoretical models can be meaningfully mapped into computational mechanisms underlying emotional behavior. Diverse approaches to computational emotion modeling, such as symbolic artificial intelligence, embodied artificial intelligence and robotics, artificial life, neural networks, and machine learning, provide different conceptual frameworks, representations and techniques to achieve these mappings. These computational models and systems address particular aspects of the roles that emotion plays in cognition and action, often disregarding their possible integration with other relevant systems, and they often stem from very different modeling traditions. In addition, many of them are guided by rather engineering concerns, lacking an appropriate understanding of relevant theories and models of human emotions that could set their grounds.

HUMAINE proposes a much-needed integrative effort in this area with the aim to shed light towards the development of sound computational models of emotions that at the same time (a) enhance the behavior of emotion-oriented systems and our interactions with them, and (b) provide feedback to emotion theorists to gain further insight in their understanding of human emotions. In this respect, the contribution of such models would be twofold. On the one hand, endowing the observable behavior of the artifact with the features of autonomy and coherence that are required to achieve long-term interactions adapted to humans. On the other hand, contributing towards a better understanding of human emotions by providing a synthetic approach (by building systems) that complements the analytic studies carried out in disciplines such as psychology and cognitive neuroscience.

1.1.2 The current conception

Our current conception is this workpackage is an elaboration of the view reflected in the TA. The focus of WP7 was identified in the TA to be a principled effort to achieve a better understanding of basic issues and open research topics regarding the involvement of emotions in cognition and action, with a view to grounding and promoting sound research into artificial emotional systems for artifacts that must interact with humans. The main task of WP 7 was further specified in D7b (page 5) as “the investigation of (computational) ‘internal’ mechanisms (emotion architectures) that allow to synthesize or generate emotions and to model their involvement in various aspects of cognition and action in emotion-oriented systems.” This task clearly necessitates an integrated effort spanning different disciplines, rather than the development of isolated engineering projects. In the TA, the need for a shared critical reflection in this area was identified regarding different aspects: key conceptual issues, assumptions and dependencies; open research problems; key research and application scenarios; analysis of needs and directions for future research and applications based on a critical analysis of existing approaches, systems and tools; evaluation methods, scenarios and tools; analysis of needs for recommendations of good practice; requirements for usability; and potential for cross-fertilization among disciplines. Our exemplar proposal has been conceived to address these aspects in a principled and tractable way.

1.2 The exemplar proposal

Following the consultation period, the exemplar proposed for WP 7 is titled *Comparative Approaches to Emotion-Oriented Architectures: Assumptions, Integration Challenges and Guidelines for Future Research*. It has four main elements within and across which our integrative effort will take place.

1.2.1 The significance of the title

This title has been chosen to stress various ideas:

- “Comparative approaches” wants to emphasize the fact that this WP acknowledges and welcomes the diversity of conceptual and computational models and frameworks that are and can be used to model emotional systems. These different approaches are not necessarily equivalent or redundant, and the task of this exemplar is to understand (in a deep sense) their scope, limits, incompatibilities and complementary aspects.
- “Comparative approaches” also wants to de-emphasize the idea of a “unified” design or model for an emotion-based architecture.

- The subtitle “Assumptions, integration challenges and guidelines for future research” stresses the nature of our common, principled integration effort in setting sound grounds that can guide future research in the area.

1.2.2 The elements of the exemplar

The proposed exemplar is divided in four main elements or sub-problems corresponding to major approaches in the conceptualization and computational modeling of emotions and their influence in cognition and action in emotion-oriented architectures. Each element will be addressed by a working group (WG):

- Element 1 (WG1): Emotion in “lower-level” cognition and action
- Element 2 (WG2): Emotion in “higher-level” cognition and action
- Element 3 (WG3): Bridging the gap between “lower-level” and “higher-level” cognition and action
- Element 4 (WG4): Emotion in Social Cognition and Interaction

Integration efforts in the investigation of assumptions, integration challenges, and guidelines for future research will take place at three levels: within working groups, across working groups within the workpackage, and across workpackages.

1.2.3 The ways in which the exemplar depends on other Workpackages

Relevant links with other workpackages and their exemplars are further developed in Section 2.5 (“Measures taken to ensure coherence with other exemplars”), and mentioned whenever deemed appropriate for different elements of the exemplar in Section 4 (“Proposed Allocation of Tasks”). In this section, we give an outline of the main types of input that WP 7 would need from other workpackages.

From WP 3:

- Overview of different theories and models of emotions in non-computational disciplines such as psychology, neuroscience, philosophy and sociology
- Guidelines for the principled design of emotion-oriented architectures
- Analysis of computational models / implementations in terms of theoretical models

From WP 4: Guidelines on emotion recognition (how to recognize signs of emotion from signals) in different approaches to emotion-oriented architectures and platforms (or “embodiments”).

From WP 5: Corpus of natural dialogues to perform linguistic analysis of user’s interactions with emotion-oriented systems.

From WP 6: Guidelines on the generation of appropriate expressive behavior.

From WP 8: Guidelines on models of persuasion and deception

From WP 9: Evaluation (of performance) will take place internally within the WP7 exemplar, but no connection with WP 9 (evaluation from the point of view of design) is anticipated at present

1.2.4 Proposed final output

Working groups and institutions will develop different types of outputs (designs, implementations, experimental and theoretical analysis, etc) as part of / as a basis for their contribution to this exemplar. However, we consider those outputs as secondary. The fact that the WP has one main task and that the exemplar is an illustration of how to undertake this task in a principled way should be reflected in the nature of the output: a single piece summarizing the investigation of the assumptions, integration challenges, and guidelines for future research issuing from our comparative approach.

To promote this type of integration, the proposed main common output arising from this exemplar is an edited book, with contributions reflecting our critical and integration efforts. The contents of this book would include:

- Theoretical chapters:
 - One general chapter on the aims, program and achievements of the exemplar (normally co-authored by the WP leader and the coordinators of the different working groups)
 - One chapter per working group addressing the results of their particular element of the exemplar from a theoretical or conceptual point of view (co-authored e.g. by the people more actively involved in each working group, or one person per institution)
- More “practical” or “technical” contributions: Various chapters on systems, designs, experimental studies, scenarios, etc, developed within the exemplar (co-authored by relevant contributors)
- Chapters about links with other workpackages

1.3 Related resources

The main HUMAINE resources complementing this deliverable are the following:

The Technical Annex expands on our early conception of the area and of (some of) the main problems that interdisciplinary integration efforts like HUMAINE need to solve to promote sound progress. It also provides a comprehensive list of references on the different scientific areas of the network, including WP 7.

Deliverable D2a summarizes discussions held at the First Plenary Meeting regarding the definition of the task of the workpackage and the nature of the exemplars.

Deliverable D7b provides a detailed review of the state of the art in the area: main achievements, key challenges and development goals, including references to key publications on different problems of the area; it describes our first proposals for exemplars, already pointing to the perceived need to cover all the elements detailed in this report, and including descriptions of the expertise that the different groups could bring towards their realization; it also contains definitions of the main terms relevant to the exemplar.

Deliverable D3c contains additional definitions of terms and provides an overview of the main theories and models of emotions from the perspective of psychology.

The portal (in particular the discussion group of wp7) contains some of the draft documents used in the preparation of this deliverable and some of the discussion (very few since most of them have taken place over normal email, telephone and in person).

2 Rationale for the exemplar proposal

The exemplar proposal represents a choice to follow a particular line of development rather than others that are possible (or might seem possible to the outsider). The key reasons for making this particular choice are as follows.

2.1 Distinctive features of the approach proposed

The main features that characterize our approach are its explicit and strong comparative and “meta-level” aspects, and the title has been appropriately chosen to stress these aspects, as explained previously (refer to Section 1.2.1, “The significance of the title”).

It is the principled integration effort that constitutes our main expected contribution to the field. This contribution is principally of a meta-level nature: even when a group decides to produce work for the exemplar that is an implemented, working system, the main contribution of that system will be its rationale – why and how it has been developed in a principled way, an understanding of its assumptions and scope, how it contributes to the advancement of the field and how it can guide future work – rather than the actual system alone, which in itself, without that rationale, would only constitute “yet another system.”

These aspects explicitly shape the overall structure of the activities of our exemplar (and of each of the elements within), which will unfold in three main stages, corresponding to the three elements of the subtitle (see Section 5 for a detailed timeline):

1. *Assumptions*: Uncovering of assumptions underlying our existing models, tools, etc. This should be completed to feed into the deliverable describing the final exemplar (Month 18, D7d) and discussed at the workshop (Month 19, D7a).
2. *Integration challenges*: Investigation of key problems, integration challenges (at different levels: cross-disciplinary, paradigmatic, conceptual, technical, etc.) and contribution towards key development goals, in particular those identified in D7b. This is the bulk of the work carried out within the exemplar and will be developed for the best part of the project (Months 19-39). It should include evaluation of the work done.
3. *Guidelines for future research*: Based on our analysis in the first stage and the evaluation and experience acquired from our practical / theoretical work in stage 2, during the last stage (Months 40-48) we will propose guidelines for future research in this area and write the final contributions to the WP 7 book.

The reasons for adopting this approach are provided in the next section.

2.2 Rationale for emphasizing this approach

The rationale underlying our approach might be better understood by examining the reasons for discarding potential alternatives. Two main alternative ways of developing an exemplar might seem natural on a first approximation:

One alternative could be the “design of an architecture for an effectively competent agent” (the idea of the actual implementation of such architecture is too unrealistic to even consider it), echoing the theoretical counterpart of this in the exemplar of WP 3 (the “blueprint”). We

discarded this idea very early in the project, and the title of our exemplar wants to de-emphasize this idea explicitly. We consider that a single, unified, global emotion architecture is neither desirable nor possible given the current state of the art (and some of us would say neither desirable nor possible in principle). The different conceptual problems and integration challenges reviewed in deliverable D7b point to the fact that such a “unified architecture” would be a misleading goal for the field that, at best, could only be done in an ad hoc, superficial and partial manner. Sound progress in the field requires at this point an effort to achieve a shared critical understanding of current achievements and needs.

The other alternative could have been to divide the task in different cognitive and behavioral “modules” (such as “perception”, “learning”, “attention”, “navigation”, etc) and investigate how emotions affect each of them individually. The impracticalities of such approach seem all too obvious: the need to make a choice of particular “modules” among the many possibilities, the blurry boundaries among many of them, the difficulty to combine results from different modules, etc. This division would have reproduced, for the study of emotion, the way in which classic AI divided intelligence in “cognitive functions”, and therefore it would have inherited the same problems encountered by classic AI, such as the difficulty in integrating (in a deep way) those “modules”, the exclusion of embodiment, problems of symbol grounding and meaning, etc. In addition, this approach would have clearly sided with one of the AI traditions only, to the exclusion of the rest. Therefore, this approach was also discarded very soon.

The approach we have opted for seems to us the best suited to meet the objectives of HUMAINE to promote the advancement of research in this area in a principled way, as this necessitates: (a) a thorough understanding of our current models and systems and the way their scope and applicability are biased as a consequence of often unnoticed or poorly understood theoretical and technical positions (their “assumption”); (b) careful thinking about the “integration challenges” at multiple levels that constitute bottlenecks hampering the advancement of research; (c) awareness within the research community of sound reflection about the current problems in the area and ways to overcome them (“guidelines”); and (d) a joint effort by researchers of different disciplines and traditions from which no relevant party is excluded, and in which the goal is not to show which approach is “right” or “best”, but to understand what each of them can and cannot offer to the global “research map” – a “comparative approach”.

2.3 Rationale for subdividing the task

From the beginning of HUMAINE, the very broad scope of this workpackage called for the need to:

- 1) Focus our efforts around a particular aspect of cognition and action, but rich enough from the point of view of the capabilities involved and the influence that emotion has on it. At the meeting of WP 7 held during the First Plenary Meeting, it was decided that we would focus on the study of the influence of emotions in decision making, since decision making involves many aspects of cognition and action and is influenced by emotion in important ways.
- 2) Approach the above problem in a tractable way, for which some division of the task was needed. Following our philosophy of a “comparative approach”, a division of the problem in independent sub-problems was not a good approach. We rather opted for dividing the task according to the main traditions towards cognition and action in AI (also related to the main traditions towards emotion, cognition and action in other disciplines such as

philosophy and psychology), which also take different views regarding the way they conceptualize and model emotion. These different traditions focus on aspects of emotions in cognition and action that are partly very different and partly “overlapping” (some of the problems involved are of the same type or about the same issue but they are approached differently). This division promotes at the same time work on a common task and a sound investigation and comparison of the contributions and perspectives of the different approaches. The exemplar has been divided into four elements, each informed by one of these main approaches. The approach and problematic of each element will be detailed in the remaining sections of this document.

2.4 Measures taken to ensure coherence across elements

The four elements somewhat constitute different “facets” of the problem. It is therefore very important to ensure their cohesion. Several measures have been taken and are planned to facilitate this, such as:

- Each element or working group (WG) has a coordinator in order to ease, from a practical point of view, coordination, exchange of information and discussion across elements. The WG coordinators will be in contact regularly (in particular over email and telephone) to this end.
- The calendar of activities has been planned at the level of the exemplar (and not of each element) and therefore its core is the same across elements: all the elements follow the same plan of activities, that each of them implements to address the particular problems considered under the approach in question.
- Joint discussions of results are programmed regularly within the plan of activities of the exemplar. This will be further reinforced by the fact that a number of participants are involved in several working groups, and this will naturally favor joint discussion not only of results, but also of problems and methods.
- The deliverables produced issue from these joint activities, and integrate the results from all the elements. This does not prevent each element from producing intermediate “internal” working documents and publications as needed to facilitate their task and disseminate results.
- The output of the exemplar is a single piece of work (a book) to which all the elements contribute. At least two chapters will explicitly take a global view of the exemplar.

2.5 Measures taken to ensure coherence with other exemplars

This exemplar presents clear links with those of other workpackages, in particular WPs 3, 4, 6, 8 and 10. Several measures have been taken and are planned to ensure coherent links with those exemplars, namely:

- Active participation in the workshops and other meetings of those WPs:
 - The workpackage leader presented and discussed cross-links of our proposed exemplar at the workshops of WPs 3 and 4, and further presentations and discussions by the WP leader or appropriate WP 7 representatives are planned for other workshops.

- The WP leader and other WP 7 representatives participated in the meetings of WP6, WP8 and WP9 that took place on 28 October 2004 in Paris.
- Similarly, the workshop of WP 7 will solicit active participation by the leaders and/or other representatives of other relevant workpackages.
- Cross-workpackage meetings will be held during the project, two have already taken place:
 - Meeting with leaders or representatives of WPs 3, 4 and 6 and the project coordinator during the workshop of WP4 in Santorini.
 - Special meeting of the PSB and other WPs representatives in October 29-30 2004 in Paris. At this meeting, the links and inter-dependencies with all the other workpackages were carefully analyzed, feeding back to the definition of the exemplar.
- Inclusion of explicit links with other WPs in various elements of the exemplar, in particular:
 - With WP3: (a) The four elements of the exemplar will provide input to the “blueprint for an affectively competent agent” element of the WP3 exemplar; (b) Explicit guidance from particular emotion theories will be used in the design of some of the computational architectures developed within our exemplar (in particular in elements 1 and 2); (c) Analysis and commentary of some of our computational models will be provided by emotion theorists from different disciplines, in particular neuroscience, psychology and philosophy.
 - With WP 4: Input (at a minimum in the form of guidelines) regarding the recognition of signs of emotion from signals has been solicited by all the elements of the exemplar. Conversely, the different elements are willing to provide feedback in the form of constraints imposed by their different approaches and platforms on the nature / interpretation / processing of signs of emotion. In addition, some kinds of models (e.g. cognitive ones, Markov-based analysis based on ethological methods) may contribute to the problem of integrating consistently the output (or input, in the case of Markov-based models) of various multimodal recognition algorithms by considering the problem of uncertainty (of non-linearity in Markov-based models) inherent in every recognition approach
 - With WP 6: One aspect of element 4 of our exemplar has been explicitly designed in collaboration with the leader of WP 6 to ensure an important link between our two exemplars: the connection between an emotion-oriented (robotic) agent architecture, which generates “emotional” behavior, and the appropriate expression of that behavior to make it meaningful in social interactions, in particular with humans.
 - With WP 8: Element 2 of our exemplar is related to the “persuasion” elements of the exemplar of this WP, in particular regarding the mutual relations between the “external” manifestations of persuasion and their use and recognition in dialogues, and the mental states underlying persuasion.

- With WP 10: Philosophical analysis of our computational emotion-oriented architectures has been solicited from, and willingly accepted by Peter Goldie and Sabine Dörner. In addition to epistemological considerations, this analysis should include ethical considerations regarding the design of emotional systems for artifacts (artifacts that “have” emotions) and their impact on the way humans perceive and interact with those artifacts.

3 Technical aspects of the program of research

The task of the exemplar is divided in four elements representing the main traditions towards cognition and action in AI (also related to the main traditions towards emotion, cognition and action in other disciplines such as philosophy and psychology), which also take different views regarding the way they conceptualize and model emotion. These different traditions focus on aspects of emotions in cognition and action that are partly very different and partly “overlapping” (some of the problems involved are of the same type or about the same issue but they are approached differently). They are the following.

3.1 Element 1: Emotion in “lower-lever” cognition and action

3.1.1 Description

This element is concerned with the investigation of the influence of emotions in cognition and action from the perspective of their embodiment. Following this view, cognition and action are inseparable – tightly coupled perception-action loops, as opposed to separable input-output elements – and intelligence can only be understood and studied in the context of “complete” embodied agents in closed-loop interaction with their physical and social environment. The focus of this approach is on the dynamics of interactions between agents and their environment and on intelligence (including emotions) seen as processes, as opposed to the “final” representational structures and actions resulting from those processes. Emotions are conceptualized in the same manner – as dynamical processes strongly rooted in the embodiment of agents and highly dependent on the dynamics of the interactions with the physical and social environment. The reason for naming this element as “lower-level” stems from the fact that “lower-level” intelligence (and emotion) is the main focus of research of Embodied AI (versus “classic” AI, the approach investigated under Element 2): intelligence (including emotion) is not solely no primarily concerned with reasoning, and many issues related to “earlier” forms of intelligence and emotion stemming from their embodied nature need to be understood before those “later” or “higher-level” forms can be investigated as related to the embodiment aspects of intelligence, i.e. in terms of “complete” agents. Therefore, this element will consider “cognitive simpler” (as “classic” AI and cognitive science have traditionally considered them) or “more primitive” (from the point of view of evolution and development) forms of emotion, cognition and action; those “simpler” forms are however very complex in many ways – hence the use of the term “lower-level” (instead of “low-level”) and conversely of the term “higher-level” (instead of “high-level”) in Element 2. Decision making is seen as action or behavior selection in this perspective: selecting what action to execute next, taking account of a combination of internal (physiological, motivational, etc) and external (stimuli) factors, in order to satisfy the current needs and survive in the environment where the agent is situated.

An important clarification is needed here: following the conceptualization of emotions as dynamic processes, this approach is aware that “lower-level” and “higher-level” aspects of

emotions and intelligence are highly intertwined in humans, and that “higher-level” aspects influence and modify “lower-level” aspects and vice versa. While the investigation of these mutual influences is undertaken under Element 3 of the exemplar, Element 1 aims to uncover the actual potential of so-called “lower-level” aspects, thereby contributing to the untangling of this Gordian knot in a bottom-up fashion.

In terms of the key problems and developmental goals described in D7b, this element addresses the following, from the perspective of emotions in embodied cognition and action:

- Key problems: Mechanisms underlying the involvement of emotions in cognition and action (5.1.1); Emotion elicitors (5.1.2); Emotions as cognitive modes (5.1.3); Emotions, value systems, motivation and action (5.1.4); Integration of different theories and models of emotions in humans and animals (5.2.1)
- Key development goals: The “origins” and grounding problem of artificial emotions (6.2.1); Untangling the “knot of cognition”: the links between emotion and intelligence (6.2.3); Measuring progress: Which are the contributions of emotions to our systems? (6.2.4)

3.1.2 Main issues to address next

Even if the “Embodied” route to AI is much younger than the “symbolic” one, this approach has by now proposed a number of Perception-Action (or “lower-level” cognition-action) architectures that are well enough motivated and understood (not only in qualitative but also in quantitative terms) to serve as a solid ground on top of which this element of the exemplar can build on. The principled investigation of emotions under this approach is however more recent, although the “embodied” perspective provides a very promising framework to investigate solutions to overcome two of the main problems recurrent in present computational models of emotions: (1) the risk of “over-attribution” of functions, capabilities and “machinery” related to emotion concepts and models (a good example being the attribution of “feelings” and conscious subjective states to artifacts); (2) the lack of “grounding” of artificial emotions, since they are often included in the system on the grounds of their meaning for the human designer and user rather than due to their meaning for the artifact in interaction with its (physical and social) environment.

Two lines of research, little developed so far, can greatly contribute to avoid those two “dangers”:

- The “emergent” approach to emotion modeling, in which behavior that an observer could consider as “emotional” but that arises from the interactions of the system with its environment, rather than from explicit “emotion components” in the architecture, can contribute to avoid the risk of “over-attribution”. It can also improve our understanding of emotional phenomena (and avoid “over-design”, i.e., the elaboration of systems unnecessarily complex) by uncovering some aspects of emotions that can be accounted for by simple cognitive and behavioral mechanisms and their interactions, without the need to postulate specific “emotion machinery”.
- Emotion grounding can be better achieved by developing computational models of emotions that take developmental and evolutionary perspectives, in which emotional systems form (or “grow”) in the course of the interactions of the artifact with its (physical and social) environment over the life time of an individual (development) or a “species” (evolution). In these cases, emotions would acquire a meaning not only for the human

designer and user, who would be able to track and understand the reasons behind those particular emotional systems, but also for the artifact itself.

Additional issues that need to be addressed under this element include:

- Different ways in which embodied/robotic models of emotions and emotion theories can collaborate to gain mutual insights towards the understanding of the involvement of emotions in (embodied) cognition and action
- Development of scenarios and quantitative methods to investigate and assess (“measure”) the influence of emotions in cognition-action in natural and artificial systems.

3.2 Element 2: Emotion in “higher-level” cognition and action

3.2.1 Description

This element investigates the influence of emotions in cognition and action from the perspective of subjective perception (as rendered by introspection and mediated by linguistic accounts) and reasoning. It largely follows the approach of “classic” or symbolic AI according to which intelligence (and emotion) can be studied independently of its embodiment and decomposed into (linearly) separable elements, namely input (“perception”), central processes (“reasoning”) and output (“action”). According to this view, intelligence and emotion are appropriately modeled by symbolic representations and symbol processing rules. It thus focuses on “higher-level” cognition-action processes as accessible to conscious awareness and describable by natural languages, i.e. in terms of discrete mental states. In AI and cognitive science and psychology, this approach has been captured in so-called “cognitive models” applied to both, reasoning processes and emotions. In this perspective, decision-making involves the elaboration (by means of reasoning) of potential plans of action leading to the attainment of explicitly-specified goals and their evaluation with respect to some explicitly provided criteria in order to select the course of action most appropriate to the attainment of the desired goal.

The main task of this element is to investigate the potential of cognitive models in representing “higher-level” processes of activation of emotions, the way in which this activation is influenced by endogenous and exogenous factors (including personality and inter-agent relationship), and the role that these models may play in human-computer interaction, with particular reference to natural language dialogs.

In terms of the key problems and developmental goals described in D7b, this element addresses the following, from the perspective of emotions in terms of mental states:

- Key problems: Mechanisms underlying the involvement of emotions in cognition and action (5.1.1); Emotion elicitors (5.1.2); Emotions as cognitive modes (5.1.3); Emotions, value systems, motivation and action (5.1.4). Integration challenges related with problems of selecting the appropriate formalisms for representing uncertainty, emotion strength, decay and mixing of emotions, and with the integration of these computational models of emotions in dialog systems (5.2.2).
- Key development goals: Untangling the “knot of cognition”: the links between emotion and intelligence (6.2.3); Measuring progress: Which are the contributions of emotions to our systems? (in particular assessing the ‘sensitivity’ of these emotion models to variation in the parameters) (6.2.4)

3.2.2 Main issues to address next

Cognitive models have a large tradition in AI and cognitive science and psychology, with numerous applications to different aspects of problem solving. Different types of cognitive models, both of reasoning tasks and of emotions, have been proposed, and some of them are understood well enough to provide a solid basis for the task undertaken here, in particular Belief-Desire-Intention (BDI) models, which consider agents as rational entities having mental attitudes that determine the agent's behavior (see definition in D7b, pages 11-12), and various appraisal-based models of emotions, in particular the so-called OCC model (see D7b, Section 4.2). These models have paid particular attention to the investigation of subjective factors that give rise to the initiation and justification of action, and the activation of emotions in terms of mental states.

Humans access emotional and rational social knowledge to guide communication, relationships, commitments, and so on. In human-human dialogs, in particular, emotions decay over time, mix, are transmitted from the speaker to the interlocutor, and vice-versa, and affect their behavior. Understanding the interlocutor's emotional state is, among the others, the starting point for planning the communicative behavior to adopt in a given context, given the speaker's communicative goals. This is particularly crucial when communication is aimed at suggesting a course of action that, for some reason, the interlocutor may find difficult to follow: typically, cease smoking or change eating habits. In this case, the amount and type of information provided must be calibrated to the attitude of the interlocutors towards their behavior: their knowledge of what a 'correct' behavior is, their belief that their behavior is (partially or fully) incorrect, their intention to change it (in part or fully), and their definition of a plan to achieve this goal.

Therefore, knowledge of the cognitive and the emotional state of the interlocutor, combined with the ability to reason about the expected emotional impact of a candidate communicative plan, allows the speaker to select the best strategy for influencing the interlocutor. An advice-giving dialog system which considers the affective aspects of the speaker-user interaction therefore needs a consistent model of the user, which extends the well known BDI approach with an emotional component (BDI&E). This model should enable the system to integrate recognition of the emotional state of the user with an *interpretation of the reasons* of this state, according to the (known or presumed) facts in its knowledge base. It should, in addition, enable it to reason about the potential *impact* of a conversational move on the mental state of the user.

Cognitive models should allow us to achieve these goals: they use principles derived from cognitive psychology to reason about the link among beliefs, values and goals, and the probability of activation of emotional states; they may also employ methods which ensure the level of expressivity which is needed to represent and handle partial and uncertain knowledge, dynamic phenomena, and variability of effects according to the context. To achieve the desired goals, the main issues that the area has to address in the immediate future are thus:

- A better understanding and development of BDI&E models, and how they compare and map to other models of emotions (e.g. appraisal-based componential models)
- Integrating into cognitive models “non-basic” emotions and “non-emotional” affective mental states that are more relevant for human-computer interaction, such as anxiety, frustration, etc.
- Linking emotion recognition (what is the present emotional state of the user?) with the interpretation of its causes, which must be inferred from the (changing) context.

- Adaptation of cognitive emotion models to the user, taking particular account of individual differences and personality.

3.3 Element 3: Bridging the gap between “lower-” and “higher-level” cognition and action

3.3.1 Description

This element of the exemplar is concerned with the investigation of potential ways of bridging the gap, with a view to integrating, “lower-level” and “higher-level” aspects of cognition and action and the role that emotions may play in it. It will do so using agents that combine both “levels”, being at the same time “embodied”/reactive and deliberative. In AI, so-called hybrid systems have for two decades attempted to “bridge the gap” by bringing together the “best of both worlds”. On the one hand, embodied systems are particularly good to act in environments in which fast decisions and actions are needed and for which “unreflected” perception-action capabilities and “pre-attentive” aspects of emotions are more adequate and effective. However, their highly reactive nature and their use of local information and interactions (as opposed to explicit and global communication) make them prone to problems of short-sightedness. On the other hand, deliberative or cognitive models are particularly good at modeling problems in terms of accounts of emotions based on e.g. “higher-level” appraisals, which allow to accommodate planning, rich expectations and action sequences extending into the future. However, they perform poorly in problems involving sensor- and environment-directed (multi-task) “processing”, as they are brittle and unreceptive to changes in the world.

In humans, emotions and other processes mediated by the so-called limbic system (or functional equivalents) seem to play an important role in linking evolutionarily older and more “reactive” brain subsystems concerned with the more immediate needs of survival (the so-called reptilian brain), and the evolutionary newer brain subsystems concerned with more cognitively costly and slower processes in the cortex (the so-called neo-mammalian brain). Taking inspiration from this mediating role of emotions, the main task of this element is to explore the roles of emotions in connecting the grounded, sensory, synchronous functioning of lower-level cognition and action to the detached, representation-based, asynchronous aspects of higher levels, e.g. via theoretical emotion constructs such as action tendencies and their mapping to concrete organized lines of activities, as well as coping and regulation mechanisms. The intended principal output of this element is *not* yet another hybrid, multi-layer architecture but a critical analysis of the main integration challenges and the potential contributions of emotions in bridging this gap.

The main contributions of this working group will be a systematic investigation of the relation between these system characteristics and architectural properties – such as modularization and structuring, synchronization, data and control flow issues – and the formulation of indications for (or against) using specific architectural components for specific emotional processes (not) to model. Clearly, this work will exploit the interfaces to the other elements of the workpackage, as detailed in the descriptions of the different threads of activity, below.

This element will investigate interactions and dependencies between different levels of reasoning and affect by using test scenarios in which both of the two levels of action selection and decision-making as well as related emotional functions are indicated and can be evaluated. For evaluation purposes, minimum requirements on scenarios (e.g., the existence of multiple conflicting goals, or worth-oriented domains) also need to be agreed upon. One

approach will involve an evolution of the TABASCO (tractable appraisal-based architecture for situated cognizes) framework developed at OFAI to integrate continuous planning and advanced behavioral components developed by members of the HW team, in which emotional states link to motivations that control multiple interactive planning and execution processes of an agent, while preserving a sound grounding in psychological theories and models provided by WP3. Behavior generation will also have to include coordinating mechanisms of coping and regulation, considering different scopes of current dynamical and situational context, aspect that will be contributed by (INESC-ID). On the other end, perceptual aspects such as attention would filter, emphasize or suppress information from the external world (IST). Contributions from partners pursuing a bottom-up approach towards motivated virtual humans (EPFL) will ensure critical assessment of ecological validity of the architectural complexity, while other working group members will contribute reflections on the rationale for high-level capacities, such as Theory-of-Mind models (USC, OFAI).

In terms of the key problems and developmental goals described in D7b, this element addresses in priority the following:

- Key problems: Mechanisms underlying the involvement of emotions in cognition and action (5.1.1); Emotions as cognitive modes (5.1.3); Emotions, value systems, motivation and action (5.1.4). Integration challenges related to Problems arising from theories and models of emotions in humans and animals (5.2.1) and to Problems with computational models, representation formalisms and implemented systems (5.2.2).
- Key development goals: Dissolving the “mind-body” problem (6.2.2.) is of central relevance to this WG, with substantial overlaps with its very workprogram; Untangling the “knot of cognition”: the links between emotion and intelligence (6.2.3); Measuring progress: Which are the contributions of emotions to our systems? (6.2.4).

3.3.2 Main issues to address next

The specific perspective taken within this element focuses on the interplay of these two “levels”, themselves to be clearly characterized in terms of:

(a) Defining properties: e.g., representational features, degree of autonomy and independence from the outside world, as with time as a reified dimension that can be manipulated independently; and

(b) Dependencies: e.g., from the nature and qualities of embedding in the world, as in virtual environments that may directly serve ready-to-use percepts.

Properties of such agent architectures regarding ecological adequacy (in terms of design parsimony and coverage of desired function) and engineering utility (in terms of desirable design and runtime qualities), as well as the qualities of the engendered responses (in terms of both, theoretical plausibility and believability) need to be investigated in designs and (partial) implementations of affective decision mechanisms for agents that are situated – i.e., the temporal dimension in particular must not be neglected.

Assessment of the adequacy of the capabilities of agents depends on the functions they should fulfill, and therefore on the environment they are embedded in. Several factors influence the degree of complexity of environments: not only the number and type of agents and other entities present and their interactions, but also the functions agents are in charge of, and the richness of their interaction with human users. These environments need to be characterized in terms of relevant properties.

Seen from the viewpoint of a single agent, beyond the physical dynamics the social dynamics present in the system needs to be considered: e.g., cooperation may by itself be a performance criterion, or access to certain resources (including e.g., the attention of the human user) may be tied to the establishment and maintenance of specific social relationships. For such purposes, the role that emotion plays in social interactions needs to be covered, establishing a link with Element 4 of this exemplar.

3.4 Element 4: Emotion in Social Cognition and Interaction

3.4.1 Description

This element of the exemplar investigates the roles of emotions in social cognition and interaction. Contrary to the previous elements, emotions, cognition and action are not modeled from the perspective of the individual agent but from the point of view of the interaction itself. Therefore, the unit of analysis is not the individual (its embodiment or its mental states) but the *relations* in groups of two or more agents. These relations might appear as a result of a wide range of “internal” and “external” factors, and can be investigated from a micro- or a macro-level perspective.

This element covers a very wide range of topics. Its goal is to work towards an understanding of the thematic area, especially of the interrelations with the other elements. The outcome of this working group should therefore be something like “map” of the thematic area, on which we can locate the places of the different sub-tasks as proposed by the members of this working group in D7b (see Section 8.4) and in Section 4.4 below. These proposals address very different points on the map, but they are far from defining a territory, let alone from covering the whole region. At this point in time, however, it will be enough of an achievement to be able to delineate the white spots on the map and to bring into relation the areas under exploration. Only then will it be possible to compare approaches.

The main outcomes of this element will be contributions to the WP 7 deliverables and book in the form of a) an overview of the area (the “map”); b) specific chapters on those areas addressed by the contributing institutions in the themes reported in Section 4.4; and c) a discussion of these specific approaches with regard to the objectives of the thematic area and to the other elements of the exemplar leading to the elaboration of guidelines for future research.

In terms of the key problems and developmental goals described in D7b, this element addresses the following in priority, from social interaction perspective:

- Key problems: Mechanisms underlying the involvement of emotions in cognition and action (5.1.1); Emotion elicitors (5.1.2); Emotions as cognitive modes (5.1.3); Emotions, value systems, motivation and action (5.1.4).
- Key development goals: Dissolving the “mind-body” problem (6.2.2); Untangling the “knot of cognition”: the link between emotion and intelligence (6.2.3).

3.4.2 Main issues to address next

- A map of the problem area defined by the goal of an affective social agent (in the broadest sense) that allows to “locate” and interrelate the numerous aspects involved, some of which are dealt with by the partners contributing to this element.
- Understanding of social dynamics from a relational (as opposed to individual-centered) perspective.
- A contribution to establishing principled and well-founded links from emotion-oriented architectures to socially meaningful expressive behavior (link with WP6), and from observed expressive behavior to interpretation (link with WP4).
- Insights into the relationship between automatic and lower-level mechanisms (e.g. automatic coordination, action selection) and higher-level cognition (e.g. goal-directed behavior, decision making) in the context of social emotions

4 Proposed allocation of tasks

4.1 Element 1: Emotion in “lower-lever” cognition and action

4.1.1 Participants

Coordinator: Lola Cañamero, UH

Participants:

- UH (30 PM Jan05-Jun06): Lola Cañamero, Rene te Boekhorst, Orlando Avila-García, Arnaud Blanchard, Humaine research fellow (expected start: January 05).
- GERG (5 PM Jan05-Jun06): David Sander, Klaus Scherer
- KCL (3 PM, tbc, Jan05-Jun06): John Taylor; Peter Goldie and Sabine Dörner.
- OFAI (1 PM Jan05-Jun06): Bernhard Jung
- ICCS-NTUA (1 PM Jan05-Jun06): Kostas Karpouzis
- CNRS STIC-SDV EPML 38 (new partner): Jacqueline Nadel, Robert Soussignan on developmental psychology; Philippe Gaussier, Pierre Andry on imitation and epigenetic robotics

4.1.2 Problems to be addressed

This element investigates the influence of emotions in various aspects of cognition and action from an embodied AI perspective, where cognition-action is regarded as tightly coupled perception-action loops. Emotions, rather than in terms of basic categories or points in a dimensional space, are modeled in terms of (patterns of) “neuro-/hormonal modulation” (that we will refer to in general as “hormonal modulation” for short) acting on the same “neural substrate” – in our case a robotic architecture. Compared to other emotion models such as “circuit models” or “adaptational models”, and as discussed more extensively in D7b, this approach to emotion modeling permits a more seamless integration of emotion and cognition-action since emotions correspond to different “functioning modes” of the same underlying “neural substrate” and associated cognitive and behavioral “functions” such as perception, attention, prioritization of motivations, behavior selection, and behavior execution. We will use autonomous robots as platforms for the work of this element of the exemplar.

Expected outputs of this element, in addition to contributions to the WP 7 deliverables, workshop and book, are publications in journals and relevant conferences (at least one per theme) that will be made available to the network.

4.1.2.1 Systematic investigation of emotional modulation of cognition and action

This theme takes an incremental, bottom-up perspective to the study of emotional modulation of cognition-behavior. Rather than taking inspiration from and trying to explicitly model theories of human emotions relative to how emotions affect different cognitive and behavioral functions, we start from a basic architecture and explore the space of potential emotional

influences on it: how “modulation” of different aspects of the architecture alters its functioning mode in different ways. Critical analysis of the results by emotion theorists (GERG and KCL) from the perspectives of different disciplines (neuroscience, psychology and philosophy) will seek to draw similarities and differences between our “artificial emotion” studies and emotions in natural systems (humans and other animals), this way assessing the plausibility of our results and at the same time providing feedback or insights into the understanding of mechanisms underlying the effects of emotions in natural systems.

For this analysis, we plan to use initially the motivated action selection architectures developed by members of the UH group (Avila-García and Cañamero), applying modulation to: (a) Perception (interoception and exteroception), i.e. altering sensor readings; (b) Action: altering motor commands / behavior execution; (c) the “links” between both, i.e. alteration of the structure of the connections. We will study how modulation of those different aspects alters (and in which ways) the resulting behavior of the robot(s). A second step would investigate the application of “hormonal modulation” to different architectures of increasing cognitive/behavioral complexity, exploring how the potential for modulation varies (increases) as a function of the complexity of the underlying “neural substrate”, i.e., more “complex” cognitive-behavioral functions would be more prone to the influence of emotions.

4.1.2.2 Principled design of “E-motivated” architecture for action selection

This study complements the previous one as it will explicitly use input and feedback from emotion (appraisal) theory (GERG) to guide some aspects of the design of the above-mentioned “modulated” and “motivated” (that we will shorten in “E-motivated”) action selection architecture, and will focus on modulation of perception (link with WP4 on recognition of signs of emotions from signals in robots).

The study of mechanisms underlying the interrelation between emotion and cognition-action does not necessarily provide insight regarding the mechanisms that elicit emotions. Whereas a model of emotions in terms of “modulation” can show how emotions can affect various aspects of cognition-action, it does not provide an answer regarding what triggers those (patterns of) modulation and why, i.e. the process of assessment of (internal or external) stimuli. Modulation of the perception of stimuli provides a mechanism to alter the “significance” or importance given to stimuli when making a decision or selecting the action to execute, and therefore to influence the selection and execution of behavior, increasing the richness of the displayed behavioral phenomena. Using appraisal theory, we will investigate how appraisal processes could be determinant for the generation of these “hormones” and the modulation of sensory and attentional processing, therefore trying to identify when and how emotional processes appear and bias the relevance given to different sensory inputs, giving in turn rise to richer and more flexible behavior, better adapted to environmental circumstances. Appraisal theory would provide input regarding the nature of stimuli and the features that are relevant for emotion activation, and guidance to extend the design of the action selection architecture to include additional mechanisms needed to process those features from the perspective of their emotional significance.

4.1.2.3 The roles of affect in (learning by) imitation

Imitation allows to create an affective "awareness" and coordination, and to alter affective bonds by allowing agents to be capable of mimicking certain aspects of another, so that synchrony of behavior may in some circumstances also lead to a synchronization of emotions. At UH, Arnaud Blanchard and Lola Cañamero are investigating the role of affect in low-level imitation using robots. By low-level imitation we understand rather automatic processes of

synchronization and coordination of behaviors that do not involve an explicit “intention” to imitate nor reasoning about the behaviour of the other agent, its mental, or affective state. An example of low-level imitation related to affect is the phenomenon of emotional contagion, in which an agent mimics automatically the expressive behaviour of another and, by doing so, enters in a similar (internal) affective state. Other roles that affect play in low-level imitation include influencing the choice of what, when, why and to whom to imitate or not to imitate based on affective bonds developed during the interaction. Collaborative efforts between roboticists and psychologists from various teams of this working group (UH, GERG, CNRS-STIC-SDV EPML 38) will define scenarios to investigate some of the relationships between affect and (learning by) imitation in a principled way, bringing also a developmental perspective to this study. Some of these scenarios will be implemented in robots, others would constitute theoretical studies that could be used to define future collaborative research projects. Some examples of topics of interest for this theme include the role of emotion as motivator in guiding imitation or learning by imitation, the use of imitation to influence the behavior of another robot by influencing its affective state, and the development of emotions in the context of imitation.

4.1.2.4 Assessment of the influences of emotion in the performance and behavior of agents

We will address the issue of how to measure “progress”, one of the key developmental goals identified in D7b, by developing ways to quantify and measure the influence of emotions in cognition-action from two perspectives: (a) from “inside” the architecture and (b) from observed behavior.

4.1.2.4.1 Measuring performance in terms of internal viability

We will analyze how emotions influence the performance of autonomous robots and affect their survival in terms of viability: the maintenance of the stability of their “internal milieu” as modeled in terms of homeostatically-controlled internal variables. Members of the UH team (Avila-García and Cañamero) have defined a number of quantitative performance indicators that measure different ways in which “good” viability is preserved. We will apply and investigate those viability-based performance indicators in our studies.

4.1.2.4.2 Analysis of behavior

Analysis of observed behavior will employ the same methods by which ethologists analyze behavior in animals. Ethologists use a suit of statistical tools (e.g., Markov modeling, cluster- and factor analysis) to reveal temporal patterns in the sequence of behavioral elements. These patterns are then used in the construction of causal explanations for various aspects of motivation-mediated behavioral phenomena such as action selection. Much less attention has been devoted so far to the explanation of emotion-mediated behavioral phenomena. Another aspect that makes our approach depart from mainstream ethological analysis is the adoption of a non-linear dynamical systems approach. The aim of the ethological analysis in this case is to uncover the factors affecting “emotions” by modeling Markov chains after systematically varying experimental conditions (spatial distribution of the resources, initial starting position of the robots, presence or absence of the competitor, inactivation of the “emotional” modulation, etc.). Initial ethological analysis of motivated and “modulated” behavior carried out by Rene te Boekhorst has already pointed to important issues:

- Working with robots in the “real world” brings to light “side effects” that would not show up or would go unnoticed in simulations
- Ethological methods pick up the manifestations of these side effects in a systematic way (although it is not guaranteed that they can also *explain* the behavioral patterns)
- Some of these “side effects” are easily (and often unnecessarily) over-interpreted by human observers as reflecting specific “drives”, “motivations” or “emotions”

This element of the exemplar brings together different conceptual traditions from various disciplines, providing a platform to investigate the links and mappings between various theories and models of emotion that can provide input to WP3, in particular:

- Links between psychology and neuroscience: appraisal theory and neuromodulation (as embodied in a “behaviour-based” robotic architecture that also incorporates a cybernetics, dynamical systems model)
- Links between natural and artificial (robotic) models

4.2 Element 2: Emotion in “higher-level” cognition and action

4.2.1 Participants

Coordinator: Fiorella de Rosis, DI-BARI.

Institutions and researchers:

- DI-BARI (5 PM Jan05-Jun06): Valeria Carofiglio, Fiorella de Rosis, Nicole Novielli, Irene Mazzotta
- France Telecom RD (1 PM Jan05-Jun06): Franck Panaget
- GERG (5 PM Jan05-Jun06): Etienne Roesch, Klaus Scherer
- QUB (1 PM Jan05-Jun06): Edelle McMahon
- UA (1 PM): Elisabeth André
- ISTC (new member): Cristiano Castelfranchi and his group have expressed their interest in participating in this element of the exemplar.

4.2.2 Problems to be addressed

This element will investigate the problems below in the context of decision making. Support can be provided by QUB in the form of a report on the effects of emotions in cognitive processes such as attention and decision making, and later on results of their empirical study of this issue using a car-driving simulator. This would permit to link the work of this element of the exemplar with studies of how emotions affect decision making in humans.

4.2.2.1 Assessment of emotion models

4.2.2.1.1 Evaluation of cognitive models of emotion activation

The idea of cognitive models is to simulate how an individual emotion or a blending of emotions may be activated by some external event. The models are ‘cognitive’ as they represent emotion activation in terms of first- and second-order beliefs, goals, beliefs about goal achievement etc. The Bari group is refining the models developed so far (event-based emotions according to the OCC theory) by evaluating them in terms of sensitivity to individual parameters and of uncertainty in the level of probability that a blending of emotions will be activated in a given condition. These models will be connected to an agent player (HapteK), to assess the appropriateness of emotional expressions in these agents (links with WPs 6 and 9), especially when the agent has to express more than one emotion.

4.2.2.1.2 Extension of cognitive models to new emotion categories

We will focus our work on those emotions which are particularly relevant for human-computer interaction, like anxiety, frustration, satisfaction, pride etc., which have received very little attention in this area. The main goal of this part of our job is to assess whether the formalism employed to represent activation of event-based emotions (dynamic belief networks) applies also to other kinds of emotions (or attitudes?). This part of the work will be done in collaboration with GERG and constitutes a strong link with WP3 (a visit from Bari to the University of Geneva has been submitted and approved in the first Call for Interchanges).

4.2.2.1.3 From BDI to BDI&E models: Analysis of the scope and potential of BDI&E models

The investigation of this issue will include, as a first step, collecting literature and reflecting on the origins of BDI models and the role they played in the multi-agent and the user modeling community (which can be made available to the network in form of a document or a webpage). As a second step, we will examine proposals and experiences on how these models may be extended to include emotions, attitudes and other affective factors (including our own experiences in this domain). We plan to compare the BDI&E models developed by various partners in WP7, for instance by defining a common scenario on which each of us might apply his/her own methodology, in order to draw some conclusion on advantages and limits of the various approaches.

To summarize and document our work on cognitive models of emotions, we will prepare a report that will contribute to the Deliverable of WP7 planned for Month 29 and to the final book, by considering the following aspects of emotion modeling:

- Discuss whether, how and to what extent cognitive models may contribute to describe the underlying (cognitive and affective) processes which create the emotion; discuss how these models may be employed in a ‘generation’ perspective and to interpret, as well, a recognized emotional state.
- Prepare a plan of experiments on how to evaluate this kind of models and, in particular, their sensitivity to variations in the parameters included in the models themselves.

- Outline a critical evaluation of the simulation programs developed by DI-BARI and by France Telecom, with a joint plan of how it may be revised and refined.

4.2.2.2 Recognition and interpretation of emotional states

The idea behind this work is that recognition of emotions in interactions is difficult, gradual and dynamic. In addition, the hypothesis is that recognition of emotions and attitudes is of not much use if it is not integrated with an interpretation of its causes (recognizing that the interlocutor is ‘angry’ or ‘distressed’ is not sufficient to generate an appropriate answer, unless one understand what produced this state). Henceforth, integration of recognition methods with cognitive models is for their interpretation of the user’s emotional states (this provides a link with WP4). Our goal, in particular, is (i) to ‘recognize’ (with some degree of uncertainty) the emotional state of the users in natural language interaction, by means of a linguistic analysis of their moves; and (ii) to ‘interpret’ this emotional state by integrating this knowledge in a model of the interlocutor’s mind, to guess about their possible causes. Recognition will be achieved by analyzing a corpus of ‘natural’ dialogs and by comparing them with a corpus of dialogs collected with a Wizard of Oz tool (connection with WP5 and WP9). Interpretation is made by ‘propagating this evidence’ in a model which represents the reasons why some linguistic expressions may be pronounced, in terms of mental state components (beliefs, goals, etc). Interpretation and recognition are seen, in these models and methods, as dynamic processes, in which the image of the user becomes clearer as far as interaction progresses. UA can support this task with their expertise regarding the links between recognition and interpretation, e.g. in the form of a document or a presentation at the WP 7 workshop, providing a link with WP4.

This aspect, together with related contributions of the Bari group to WPs 8 and 9, will be integrated in an affective dialog simulator that this group has been developing for the last 4 years, ensuring the cohesion of results with these other workpackages.

4.2.2.3 Application framework

As we said in the beginning of this Section, the main application scenarios for our models are dialog systems. The groups in Bari and France Telecom RD have extended experience in these systems and will be particularly active contributing to this aspect.

In the context of (phone-based) spoken applications (such as call centers, information-retrieval services, e-commerce), we will consider the following issues:

- 1) Define a model in which the behavior of the system takes into account the user’s emotional state, or more precisely, the main emotions that can be observed in such applications (e.g., angry, disappointed...)
- 2) Illustrate this model on a Telecom application
- 3) Evaluate the user’s satisfaction according to emotion processing.

We will integrate a small set of user’s emotions (1 or 2 such as angry, disappointment...) into a BDI model, and show how a spoken dialogue system (applied to a Telecom application such as a call-center or information retrieval services) can adapt its communicative behavior. The expected results could be an overall framework of such a spoken dialogue system and a set of formal rules expressing communicative adaptation to user’s emotion, which could be made available to the network in the form of a document at the end of 2005 or early 2006.

If resources are found from France Telecom's internal projects, additional work could try to illustrate the use of this BDI&E model on a telecom application and realize a user's test, which could give rise to a report on the evaluation of users' satisfaction when interacting with a spoken dialogue system with and without adaptation to users' emotions.

4.3 Element 3: Bridging the gap between “lower-” and “higher-level” cognition and action

4.3.1 Participants

Coordinator: Paolo Petta, OFAI

Institutions and researchers:

- OFAI (15 PM Jan05-Jun06): Paolo Petta, Stefan Rank
- HW (18 PM Jan05-Jun06): Ruth Aylett, Pablo Lucas dos Anjos
- INESC-ID (9 PM Jan05-Jun06): João Dias
- IST (3 PM Jan05-Jun06): Carlos Martinho
- EPFL (9 PM Jan05-Jun06): Etienne de Sevin, Daniel Thalmann
- USC (new member): Stacy Marsella, Jonathan Gratch

As all members of the working group bring in working experience with control architectures for emotional agents, joint contributions are planned on all Threads of activity of this element.

The focus of contributions of Pablo Lucas do Anjos (HW) and João Dias (INESC-ID) will come from a background in behavior-based robotics and continuous planning; contributions by Etienne de Sevin (EPFL) will focus on a bottom-up view, extending a reactive motivational architecture towards more sophisticated cognitive capabilities and related emotional functions; contributions from Stefan Rank, Paolo Petta (OFAI), Carlos Martinho (IST), Jonathan Gratch and Stacy Marsella (USC) will provide inputs on conceptual clarifications, mappings between theory and engineering domain, and the scenario description language.

4.3.2 Problems to be addressed

In the following, we describe the task to be carried out in this element in terms of five main threads of activities. While there exists a logical relation between these threads – reflected in the sequence they are given – in practice work will necessarily be carried out in a less strict order and involving iterations within and across threads. These shortcuts and loops are influenced by bootstrapping necessities and global deadlines, such as the WP7 workshop, but may be also triggered by insights gained from the actual working materials contributed and uncovered. Even so, the importance of the structure given goes beyond a mere documentation of kinds of activity to be pursued, as it will be instrumental in facilitating detailed scheduling of specific subtasks to be assigned to smaller scale co-operations between WG members and serve as constant reference to ensure the coherence and avoid as far as possible idiosyncrasies of the results obtained.

4.3.2.1 Thread A: Conceptual clarifications

This thread covers the necessary terminological and functional grounding of the work in the WG, the establishing and maintenance of a shared understanding across the domains of cognitive science, emotion theory, and computational modeling, including the removal of the scare quotes that so far surround the terms “lower-level” and “higher-level” cognition and action. In this cross-WP dialog with WP3, different notions of levels of cognition from a cognitive science point of view will be contributed. Advantage will be taken of collaborations with the other workgroups; however, the scope of the concepts at the interfaces to these other groups will likely need to be extended and adjusted: the notion of lower-level cognition and action, towards richer and more varied contexts than those focused on in Element 1 and including in particular the nature of lower-level cognition and action in virtual environments; the concept of higher-level cognition and action with a focus on what is accepted from and propagated downwards to the lower level, in the transitions between e.g., situatedness (being constrained by grounding and environmental chronicity) and detachedness (constrained primarily by properties of representations and tractability).

Informed by these first clarifications, the relation of emotion theories to lower-level and higher-level cognition and action will be assessed (with possible further contributions to the taxonomy), leading to the identification of aspects of theories of pertinence to the role of emotions in bridging the gap.

The final output of this thread is a characterization of emotion theories in terms of the bridging functions covered, along with the related notions of higher-level and lower-level cognition and action. This thread involves close interaction with the other elements of the exemplar. It will provide input to the WP 7 report on “assumptions”.

4.3.2.2 Thread B: Operationalization of theory

This thread covers the mapping from (parts of) natural emotion theories pertinent to the scope of Element 3 (as identified in Thread A) to design specifications of complete synthetic emotion designs. For each of the theories, application domain related requirements can be derived, that characterize which scenarios require and warrant the providing of these specific emotional functionalities bridging this/these gap(s) between the respective specific kinds of low and high-level cognition and action.

With reference to the integration challenge of Problems arising from theories and models of emotions in humans and animals (5.2.1 in D7b), scope of the theory and completeness of the mapping into the synthetic design needs to be ascertained. Further requirements may then arise when analyzing the derived synthetic design, by uncovering tacit assumptions, and identifying all components the synthetic model presupposes and connects to, in terms of theoretical constructs (both higher-level and lower-level, e.g., memory, standards and values, properties of sensors and effectors) as well as environmental properties.

Finally, deviations from natural theory due to engineering constraints need to be assessed, including incoherencies and details that were omitted on the path from theory to implementation.

The final outputs of this thread will be specifications of synthetic emotion designs, covering functionalities and limitations, environmental requirements and matching scenarios. This thread also involves close interaction with the other elements of the exemplar. It will also provide input to the WP 7 report on “assumptions”.

4.3.2.3 Thread C: Scenario definitions and evaluation of existing systems and designs

This thread collects early efforts towards the definition of relevant scenarios as well as a critical assessment of current achievements and issues. To ensure uniformity and comparability of scenario characterizations, a scenario description language defining relevant dimensions and attribute sets will be employed. Scenario descriptions relate environmental aspects (including e.g. typical episodes) to features of higher-level and lower-level cognition and action.

This review work may also consider systems that cover only one out of the “higher” and “lower” levels of cognition and action, in the event of failing to identify a sufficient number of existing systems covering both levels. Out of these systems, successes in capturing of emotional functionality bridging lower- and higher-level cognition and action will be assessed along with the identification of relevant deployment scenarios. At the same time, incoherencies and omissions on the path from theory to implementation and cases of deployment in inappropriate contexts will also be documented.

Based on these findings, positive examples will be further characterized in terms of: quality of design and implementation (to which extent gratuitous dependencies on design elements or implementation details are avoided); scope (deriving a characterization of their niche spaces and assessing their applicability to different scenarios); and potential for extension (in particular for single-level designs). Instances of existing systems failing on the different aspects investigated will be utilized to derive additional design criteria and identify theoretical issues worth documenting.

The final output of this thread will consist in the definition of scenarios to be employed to compare and assess the scope of different models and systems, and the characterization of capabilities and development potential of existing systems, along with derived contributions to theoretical and practical design guidelines. These results will make contributions to the “assumptions” and “guidelines” reports of the exemplar.

4.3.2.4 Thread D: Integration challenges

This element of the exemplar faces specific and important integration challenges related to the very nature of its task. Integration challenges particular to this element will be derived by relating identified successes and potentials of different systems along different analytic dimensions: theories; designs/models; scenarios. This will support the formulation of criteria of practical relevance for the selection of theories, designs, implementation methods, and scenarios. These findings may undergo some limited empirical validation, by carrying out implementation work wherein the actual utility of the criteria and guidelines for the realization of systems can be verified, and by subsequent application of the evaluation procedures defined in Thread C to the resulting systems.

Both integration challenges put down in the deliverable D7b, Problems arising from theories and models of emotions in humans and animals (5.2.1) and Problems with computational models, representation formalisms and implemented systems (5.2.2) are directly addressed in Threads A and B of this element. In the present context, the notions of higher-level and lower-level cognition and action will undergo a particularly careful scrutiny, with the goal of providing a differentiated view on what has been traditionally assigned to different (“higher” and “lower”) levels in layered architectures (e.g. behavior-based systems) on the one hand, and actual functionality of “higher” and “lower” cognition according to current neurophysiological and psychological insights. Collaboration with theory (WP3) shall ensure that the efforts of this workgroup to challenge existing affective architectures are carried out

from a sound understanding of theoretical concepts, and enable it to comparatively explore the space of possible architectures by relating it to theory. These efforts are aimed not only at providing guidelines on future work on architectures but also at informing a shared understanding of theory by pointing out missing details and incoherence on the path from theory to implementation.

The particular integration challenges to focus on from month 19 onwards will be decided during the “uncovering assumptions” stage of the exemplar, as they necessitate input from Threads A to C and feedback from other elements of the exemplar. The integration challenges chosen will explore how emotions can help to bridge the gap between “lower-level” and “higher-level” cognition and action (or, in terms of computational architectures, between “reactive” and “deliberative” layers), and will have to deal with issues such as management of timing and prioritization between both “levels”, how to achieve “mappings” between the contents of these “levels”, or how to translate their functionalities (i.e. how emotions can contribute to solve those issues), to give some examples.

The final outputs of this thread are a comparative review of the state of the art, comprising the characterization of achievements in terms of functionalities provided; successful design patterns; and scenarios covered, and an analysis that distinguishes between current and inherent limitations of approaches surveyed, in terms of theories considered; functionalities of emotions selected and mapped; architectural designs; and scenarios still or necessarily out of reach. They will provide input for the WP 7 deliverables planned for months 30 and 39.

4.3.2.5 Thread E Guidelines for future research directions for principled approaches

This final thread will start only after the period of the JPA currently defined (project months 12-30). Here, a selection process will identify elements from the evaluation and comprehensive synthesis relating all analytic dimensions (scenarios; designs/models; full theory); theoretical and practical design guidelines; potential contributions from emotion psychology theory not considered in engineering contexts so far; and views from cognitive science on “lower-level” and “higher-level” cognition and action considered to be relevance.

The output of this thread will be contributions to the WP 7 book and the handbook covering these elements identified.

4.4 Element 4: Emotion in Social Cognition and Interaction

4.4.1 Participants

Coordinator: Sabine Payr, OFAI

Institutions and researchers:

- OFAI (8 PM Jan05-Jun06): Sabine Payr, Paolo Petta
- MIRALab (14 PM Jan05-Jun06): Lionel Egger, Arjan Egges, HyungSeok Kim
- UH (8 PM Jan05-Jun06): Arnaud Blanchard, Lola Cañamero, Rene te Boekhorst, Susan Atwood
- DIST (5 PM Jan05-Jun06): Gualterio Volpe, Antonio Camurri, Barbara Mazzarino, Ginevra Castellano

- Paris8 (4 PM Jan05-Jun06): Chrystopher Peters, Catherine Pelachaud
- DI-BARI (3 PM Jan05-Jun06): Fiorella de Rosis, Nicole Novielli
- DFKI (2 PM, tbc Jan05-Jun06): Patrick Gebhard
- ISTC (new partner): Isabella Poggi, Maria Miceli

4.4.2 Problems to be addressed

The broad scope of this element of the exemplar, its strong interrelations with the other elements, and the poorer awareness of the theories underlying the involvement of emotions in social cognition and interaction call for a division of the task of this element into conceptual clarification and specific thematic foci. These activities will be carried out in parallel and can be seen as orthogonal, providing feedback to one another. The issues addressed under “conceptual clarification” will be applied to the various thematic foci proposed.

4.4.2.1 Conceptual clarification

An initial sketch of the main theoretical issues addressed in this task will inform the investigation of assumptions that will be carried out initially and the definition of the integration challenges to be addressed later on in the project. We propose to address in priority the three following issues:

4.4.2.1.1 The “map” of emotions in social cognition and interaction

Work on emotion in social psychology and sociology that can be used to draw the “map” is mostly recent and, as is to be expected, far from undisputed and “received knowledge”. There are three lines along which emotions in social life are conceived: (a) emotion as socially responsive, (b) emotion as socially shared and regulated, and (c) emotion as socially constituted. Clearly, these lines are listed here in order of increasing radicalism, with the third one as a rather novel and not yet very well elaborated direction. While (a) fits well into the individualistic and reactive approach to emotions predominant in psychology (and in HUMAINE), we think that it is fruitful to pursue here an approach that corresponds to the more radical version of (b) which says, roughly, that we are socialized into our ideas about emotion and into our understanding of how particular situations link up to emotions.

The basic assumption is that there is no such thing as an emotionally neutral situation or interaction. Objects, activities, social roles etc., are attributed “affective meanings” which, as a whole, make up culture. On the surface, this standpoint may appear as contradicting the predominant “individualistic” approach taken in emotion theories and models. On a deeper level, it is of course not contradictory, but it makes the field more complex if we have to take into account the ongoing interdependency between the world and the individual. We think that it is necessary to first understand the complexity before deciding where it can be reduced in computational models. Without a deeper understanding, we might never know why we fail.

Insights into emotions in the social sciences are gained mainly from experimentation and observation, i.e., from a third-person perspective. Models, however, are built from the inside out, that is, a first-person perspective. The translation of observed and observable emotional behavior to models that generate that behavior (emotion synthesis) is a non-trivial problem once one tries to take into account its numerous, often conflicting, intrinsic and extrinsic sources.

The particular “points in the map” that we will investigate in detail are described under “thematic foci”, and the study of their place in the global map will take into account a clarification effort regarding, in particular, the two issues below.

4.4.2.1.2 Emotions as cognitive modes in social interaction

The approach to emotions in social cognition and interaction proposed here views emotions as pervasive, and directly addresses the key problems and questions outlined in section 5.1.3 of D7b. From the point of view of social interaction, emotions have a continuous and global influence on our relation with the world. There is not a special “emotional” cognitive mode, contrasted with a “mode-less” normal or neutral state, but a permanent modulation of an agent’s being in the world, resulting from sources as divergent as knowledge, history, value systems, motivations, roles, situatedness, goals, and the (inter)action itself (cf. D7b, 5.1.4). This view has a direct link with the conceptualization of emotions as patterns of neuromodulations adopted in Element 1 of the exemplar. The sources considered here are, however, themselves socially constructed, so that the emotional-cognitive mode does not vary arbitrarily from agent to agent (including humans). Therefore, while a social cognitive viewpoint adds complexity to emotion synthesis, it can also constrain interpretation of emotions.

4.4.2.1.3 The “lower-level“ and “higher-level” of emotion in social cognition and interaction

What is distinguished in this exemplar as “lower-level” and “high-level” cognition are both relevant in social interaction. The work being proposed in the area of “attention” or “contagion” is, for example, clearly situated on the “lower” end of emotional/cognitive processes. The aspects of “empathy”, “facework” or “role behaviour” can be attributed to the higher-level functions. But this distinction is far from established for all the processes in interaction. Some work, for instance, has been done on suppression of emotion expression. At first sight, this would be clearly the result of conscious decision-making. However, Hochschild, in her seminal work on emotion management, has defined three different levels on which “emotion management” (which covers both suppression and generation of emotions) is done: “surface acting”, which coincides with a conscious decision-making approach, “deep acting”, whereby persons learn to recall experiences that help them to convincingly “act out” the emotion that is called for in the current situation, or “identification”, where persons learn to live their role model to such a degree that the desirable emotion is genuine and the undesirable emotion is not suppressed, but truly absent. This is only one example of the challenges facing this working group and that permeate both of its thematic foci.

4.4.2.2 **Thematic foci**

The “points on the map” that this working group proposes to address are the following:

4.4.2.2.1 Emotion in social dynamics

Under this theme the work of the exemplar is to achieve a critical understanding of (selected) key mechanisms and key factors underlying the involvement of emotions in the development, structure and dynamics of social groups. The theoretical study will be grounded by software simulations of different aspects of social dynamics already developed and to be developed by members of this working group, in particular (see D7b, Section 8.4 for extended descriptions):

- The roles of moods and appraisal in social interaction, using the multi-agent (multi-ECA) environment proposed by DFKI (Patrick Gebhard). Moods are longer lasting affective states that affect long term behavior. The information regarding the affective state (mood) of each single emotional character and the user can be used to compute an overall mood (moral) vector of the group, a social integrity vector, mood extremes, and subgroups with the same mood configuration. This would enable a new level of interaction design, such as problem coping strategies, and motivation strategies for learning environments.
- The supporting role of emotions for the sustenance of social norms, and hence their impact in the formal groups and interactions, based on the multi-agent simulation studies carried at OFAI (Petta & Staller). The function of emotions in bridging the micro-macro gap is modeled in terms of the mutual support of norms and emotions.
- The role of environmental elements and simple local affect-based interactions in the emergence of informal groups with varying organization and dynamics (e.g., hierarchical/dominance based versus peer/friendship based), using Artificial Life simulation studies carried at UH (Susan Attwood, Lola Cañamero, Rene te Boekhorst)
- Other topics such as trust could be contributed by Cristiano Castelfranchi and his group at ISTC (new partner); this is currently under discussion and subject to availability of resources.

4.4.2.2.2 Guidelines for the a socially meaningful emotionally-competent agent

To be emotionally competent in the context of social interaction in a way that is meaningful to humans, an agent must integrate an emotion-oriented architecture and at the same time express emotions through its behavior in a socially meaningful way. Under this theme, we propose to elaborate guidelines for a principled integration of an emotion-oriented architecture and synthesis of expressive behavior (link with WP 6) so that the affective state can be appropriately recognized from its behavioral manifestations (link with WP 4). Due to the complexity of each of these two aspects, they are usually developed independently, and this is the case in works with robots, ECAs, virtual humans, and other software simulations. Even within HUMAINE, these aspects have been allocated to different Workpackages (WP7 and WP6, respectively), due to the need to divide problems so that they can be addressed in a tractable way. Under this theme, the exemplar will investigate how this important gap could be bridged in a principled way, making a natural link with WP6. For this, we will use different platforms and types of “embodiments” in order to avoid producing too narrow a solution that would be biased towards one of them. To ensure the feasibility of this task, we will focus on a particular type of interaction that involves the “sharing” of the affective state by two agents. This “sharing” can take place without conscious awareness, as in the case of contagion and other rather automatic synchronization and “low-level” imitation mechanisms, or it can involve conscious awareness, as in (“higher-level”, “cognitive”) empathy. We will use the following elements:

- The emotion-oriented robotic architecture for “low-level imitation” and contagion being developed at UH (Arnaud Blanchard, Lola Cañamero). By low-level imitation we understand rather automatic processes of synchronization and coordination of behaviors that do not involve an explicit “intention” to imitate nor reasoning about the behaviour of the other agent, its mental, or affective state. An example of low-level imitation related to affect is the phenomenon of emotional contagion, in which an agent mimics automatically the expressive behaviour of another and, by doing so, enters in a similar (internal) affective state. Other roles that affect play in low-level imitation include influencing the

decision making regarding what, when, why and to who to imitate or not to imitate based on affective bonds developed during the interaction.

- “Higher-level” or more “cognitive” forms of this family of phenomena, usually known under the concept of “empathy”. DI-BARI, in cooperation with Isabella Poggi and Maria Miceli, (ISTC) propose to contribute the following work: formalization of the types and forms of empathy transmission; discussion of the types of empathy which may be transmitted from human users to agents (using ECAs) and how they may be ‘recognized’; discussion of how the agent (ECA) should respond to an ‘offer of empathy’ by the user; discussion of whether an effect of empathy induction may be hypothesized from ECAs to the user. In addition to contributing to WP7 reports on the “integration challenges”, this work will establish a link with WP9 in developing a set of design and evaluation criteria regarding the design of how to design ‘empathic’ ECAs.
- Social attention and memory mechanisms for early social interaction developed at Paris8 (Christopher Peters, Catherine Pelachaud), in the context of low-level imitation / contagion / empathy. Attention models have been used as a basis for providing fast and automatic behaviors suitable to animated agents in specific situations, ranging from virtual helicopter pilots in combat environments to conversational agents in social settings, or for modeling more general attention behaviors during free viewing conditions in virtual environments. Locations looked at by an agent may be based on current goals or may elicit the agent’s attention in other ways. Shifts of attention may result in the production of visible phenomenon that are of significance to others in the environment as indicators of goals and intention, for example, gaze and orienting behaviors. Therefore, some form of social attention mechanism is desirable for allowing particular consideration of the attention behaviors of others, i.e. ‘attention paid to attention’. This sort of mechanism may serve as a basis for the formulation of simplified agent ‘theory of mind’ models.
- Generation of socially meaningful emotional (expressive) behavior (standards for translating the above-mentioned cognitive-emotional process into face and body behavior, definition of basic “units” that control integrated face and body behavior), taking advantage of the real-time animation systems developed at MIRALab (Arjan Egges), in particular virtual humans, that result in realistic facial and body animations (see D7b, section 8.4 for an extended description).
- The use of gesture and movement in the expression of emotion in low-level imitation (synchronization, contagion) tasks based on the studies on interactive artistic performances carried at DIST (Antonio Camurri, Gualterio Volpe)
- Requirement analysis for an affective social agent and its constraints/affordances for emotion synthesis and emotion interpretation, in the form of interdisciplinary analytical work on the basis of prior studies at OFAI (Sabine Payr, Paolo Petta)

5 Provisional timeline

The calendar of activities of the exemplar follows the 3 stages outlined in Section 2.1 (“Distinctive features of the approach proposed”) and stressed in the subtitle of the exemplar. The calendar given below does not include other activities of the network (such as plenary meetings, summer schools, etc), for which we refer to the Technical Annex.

1) Stage 1: critical analysis of the current state of the art, needs, and assumptions.

- Months 1 to 11: work during these months consisted in iterations to identify key problems, development goals, expertise available within the network, and needs in order to come up with a sensible exemplar proposal for the thematic area.
- Months 12 to 18: the definition of the exemplar will be further detailed and consolidated. A key ingredient of this task will be the investigation and “uncovering” of the assumptions underlying the approaches and models of the different elements of the exemplar. One of the major obstacles preventing the integration of (theoretical or practical) work, the understanding of how the different elements of the global “map” relate, and ultimately the sound advancement of research in the field, is the poor understanding of the theoretical underpinnings and the scope and applicability of our models. The analysis of these assumptions is therefore a prerequisite to the principled understanding and investigation of the integration challenges and guidelines for future research developed during subsequent stages. It will lead to the selection of the first “integration challenge” that each element will develop during the first phase of the following step.
- Month 18: Final specification of the exemplar in deliverable D7d, grounded in the analysis of assumptions and defined to tackle problems and issues identified in this critical analysis.

2) Stage 2: Investigation of key problems, integration challenges and contribution towards key development goals

- Month 19: participation in WP7 workshop and contribution to the workshop proceedings (deliverable D7a). The main aim of the workshop will be to present the “assumptions” found in the previous critical analysis and the work proposed within each element in terms of the first set of “integration challenges” within and across elements, and with other WPs. Ideally we would also start to think about the type of guidelines and final output that we might produce by the end of the project.
- Months 19 to 30: First integration challenge. During months 19-29 each element will develop theoretical and practical work (such as designs, implementations, experimental studies, scenarios, comparative analyses, evaluations, etc., plus the integration of the different aspects) around a main problem or “integration challenge”. Month 30 will be devoted to the common discussion of the results of the different elements. This will be done in a “meeting” (which might take the form of programmed email and portal exchanges, a series of telephone conferences, an actual physical meeting, or a combination of them) and will give rise to a deliverable (D7e).
- Months 31 to 39: Second integration challenge. During months 31-38 each element will develop theoretical and practical work (such as designs, implementations, experimental studies, scenarios, comparative analyses, evaluations, etc., plus the integration of the different aspects) around a main problem or “integration challenge”. This could be the further development of the previous challenge or of a new one, depending on the

dynamics of the work and the results obtained. Month 39 will be devoted to the common discussion of the results of the different elements. This will be done in a “meeting” (which might take the form of programmed email and portal exchanges, a series of telephone conferences, an actual physical meeting, or a combination of them) and will give rise to a deliverable (D7f).

3) Stage 3: Guidelines for future research and conclusions

- Month 40: the bulk of the “integration” work should be completed by now although some work might continue. Proposals for chapters and abstracts should be sent to the WP leader and groups should start working on the chapters. During Months 40 to 42 each working group should work on guidelines for future research; guidelines will be included as part of the chapters for the book. Depending on time constraints, we might decide to collect those guidelines in deliverable (D7g) as well, to be decided later on during the elaboration of the corresponding new JPA.
- Month 43: drafts of chapters must be sent to the WP leader.
- Month 48: end of exemplar and draft of book ready to be sent to the publisher

Proposed exemplar (months 13-48)						
Workpackage number :	7	Start date or starting event:			month 13	
Exemplar:	<i>Comparative Approaches to Emotion-Oriented Architectures: Assumptions, Integration Challenges and Guidelines for Future Research</i>					
Participant id:	UH	OFAI	DI-BARI	HW	UNIGE-GERG	EPFL
Person-months (M 13-30 only):	38	24	8	18	10	9
Participant id:	INESC-ID	IST	DIST	Paris8	UNIGE-MIRALab	KCL
Person-months (M 13-30 only):	9	3	5	4	14	3 (tbc)
Participant id:	ICCS-NTUA	UA	DFKI	France RD	Telecom	
Person-months (M 13-30 only) :	1	1	2 (tbc)	1		
New members interested:	CNRS STIC-SDV EPML 38	USC	ISTC			
Person-months (M 13-30 only):	tbd	tbd	tbd			

Objectives	<ol style="list-style-type: none"> 1. Achieve a shared understanding of basic issues and open research topics 2. Set the grounds to advance the state of the art
-------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

3. Disseminate the achieved research results and conclusions

Elements	<ol style="list-style-type: none"> 1. Emotion in “lower-level” cognition and action 2. Emotion in “higher-level” cognition and action 3. Emotion in bridging the gap between “lower-level” and “higher-level” cognition and action 4. Emotion in social cognition and interaction
-----------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Description of work

Task 7b.1 Uncovering underlying assumptions

<p>A key ingredient of the exemplar is the investigation and “uncovering” of the assumptions underlying the approaches and models of the different elements of the exemplar. One of the major obstacles preventing the integration of (theoretical or practical) work, the understanding of how the different elements of the global “map” relate, and ultimately the sound advancement of research in the field, is the poor understanding of the theoretical underpinnings and the scope and applicability of our models. The analysis of these assumptions is therefore a prerequisite to the principled understanding and investigation of the integration challenges and guidelines for future research developed during subsequent stages.</p>

Task 7b.2 Integration challenges

<p>Each element of the exemplar will develop theoretical and practical work (such as designs, implementations, experimental studies, scenarios, comparative analyses, evaluations, etc., plus the integration of the different aspects) around a main problem or “integration challenge”. Each element will carry work around two “integration challenges”, one during months 19-30, the other during months 31-39.</p>

Task 7b.3 Elaboration of guidelines for future research an dissemination of results

<p>The bulk of the “integration” work should be completed by month 40 although some work might continue. Proposals for chapters and abstracts will be sent to the WP leader and groups will start working on the chapters. During Months 40 to 42 each working group should work on guidelines for future research; guidelines will be included as part of the chapters for the book. Depending on time constraints, we might decide to collect those guidelines in deliverable (D7h) as well, to be decided later on during the elaboration of the corresponding new JPA.</p>

Deliverables

D 7a (month 19): Workshop proceedings

D 7d (month 18): Final proposal for exemplar

D 7e (month 30): Report on the first set of integration challenges

D 7f (month 39): Report on the second set of integration challenges

D 7g (month 43; still to be decided): Report on guidelines for future research

Milestones and expected result

M7-3 (month 18): Decision on selection of exemplar(s) for joint research

M7-4 (month 19): Presentation of results on “assumptions” and proposals for the first set of “integration challenges” at the WP 7 workshop

M7-5 (month 29): Results on “first integration challenge” by each element of the exemplar to prepare D7e

M7-6 (month 38): Results on “second integration challenge” by each element of the exemplar to prepare D7f

M7-7 (month 43): Proposals for guidelines for future research and for contributions to WP 7 book (draft chapters)

M7-8 (month 48): Draft of WP 7 book

6 Conclusion and way Forward

This report presents our proposed potential exemplar for the area of Emotion in Cognition and Action and explains our reasons for following the proposed approach. The proposal will be refined until the final proposal is presented in Month 18 (deliverable D7d). At the same time, as a necessary step in the exemplar, we will analyze the assumptions underlying our models and systems and that determine their scope and applicability, and set constraints regarding their potential for integration with other models and systems.

The workshop of this workpackage will be held in month 19, and will serve as a forum to discuss the critical analysis of the assumptions and the first set of “integration challenges” that the different elements of the exemplar will undertake as part of the new JPA. Shortly after the beginning of the new year, the WP leader will start a consultation phase with other members of the workpackage and of the network to elaborate a first draft of the potential contributions to the workshop.