AIED 2009: 14th International Conference on Artificial Intelligence in Education

Workshops Proceedings

Editors and Co-Chairs:

Scotty D. Craig
University of Memphis, USA

Darina Dicheva
Winston-Salem State University, USA

July 6-7th, 2009
Brighton, UK
Preface

The supplementary proceedings of the workshops held in conjunction with AIED 2009, the fourteen International Conference on Artificial Intelligence in Education, July 6-7, 2009, Brighton, UK, are organized as a set of volumes - a separate one for each workshop.

The set contains the proceedings of the following workshops:

- **Volume 1: The 2nd Workshop on Question Generation**  
  *Co-chairs*: Vasile Rus & James Lester. *University of Memphis, USA & North Carolina State University, USA.*  

- **Volume 2: SWEL'09: Ontologies and Social Semantic Web for Intelligent Educational Systems**  
  *Co-chairs*: Niels Pinkwart, Darina Dicheva & Riichiro Mizoguchi. *Clausthal University of Technology, Germany; Winston-Salem State University, USA & University of Osaka, Japan.*  
  [http://compsci.wssu.edu/iis/swel/SWEL09/index.html](http://compsci.wssu.edu/iis/swel/SWEL09/index.html)

- **Volume 3: Intelligent Educational Games**  
  *Co-chairs*: H. Chad Lane, Amy Ogan & Valerie Shute. *University of Southern California, USA; Carnegie Mellon University, USA & Florida State University, USA.*  
  [http://projects.ict.usc.edu/aied09-edgames/](http://projects.ict.usc.edu/aied09-edgames/)

- **Volume 4: Scalability Issues in AIED**  
  *Co-chairs*: Lewis Johnson & Kurt VanLehn. *Alelo, Inc., USA & Arizona State University, USA.*  

- **Volume 5: Closing the Affective Loop in Intelligent Learning Environments**  
  [http://aspire.cosc.canterbury.ac.nz/AffectLoop.html](http://aspire.cosc.canterbury.ac.nz/AffectLoop.html)

  *Co-chairs*: Emmanuel G. Blanchard, H. Chad Lane & Danièle Allard. *McGill University, Canada; University of Southern California, USA & Dalhousie University, Canada.*  
While the main conference program presents an overview of the latest mature work in the field, the AIED2009 workshops are designed to provide an opportunity for in-depth discussion of current and emerging topics of interest to the AIED community. The workshops are intended to provide an informal interactive setting for participants to address current technical and research issues related to the area of Artificial Intelligence in Education and to present, discuss, and explore their new ideas and work in progress.

All workshop papers have been reviewed by committees of leading international researchers. We would like to thank each of the workshop organizers, including the program committees and additional reviewers for their efforts in the preparation and organization of the workshops.

July, 2009

Scotty D. Craig and Darina Dicheva
The Workshop on Closing the Affective Loop in Intelligent Learning Environments

Workshop Co-Chairs:

Cristina Conati  
*University of British Columbia, Canada*

Antonija Mitrovic  
*University of Canterbury, New Zealand*

http://aspire.cosc.canterbury.ac.nz/AffectLoop.html
Preface

Research on affective intelligent learning environments (ILE), i.e. ILE that include affective elements in the interaction with the student, has become increasingly prominent in the past few years, due to two main reasons. Firstly, there is growing evidence of correlations between affect and learning, fostering the belief that recognizing and responding to student's affect can improve the effectiveness of pedagogical interactions. Secondly, advances in affect recognition make it now feasible to devise interactive tools that can be aware of a user's affective states, and respond accordingly.

The workshop is designed to continue the discussion started during the previous related workshops at ITS and AIED, with a specific focus on how to close what we will call the affective loop, i.e., that ensemble of four phases that together allow for the principled addition of affective elements to an ILE: (i) design the environment so that it can elicit affective states favorable to learning; (ii) recognition/modeling of relevant user's states; (iii) selection of appropriate system responses; and (iv) synthesis of the appropriate affective expressions.

Of the four phases that form the affective loop, phase (ii) and (iv) have seen the most advances in recent years. What we are still lacking are theories on how to use affect in pedagogical interactions, and strong evidence that taking affect into account when interacting with students can actually improve learning. There are some qualitative accounts of how affect and learning may interact, as well as some evidence that certain affective states are correlated with learning, but these findings are not sufficient to devise computational models that can be used to influence the affect-sensitive behaviors of computer-based tutors. Similarly, we have initial results on the effect of affect-based ILE, but they mostly report qualitative measures of student satisfaction rather than quantitative measures of learning.

Thus, the main objective of this workshop is to bring researchers in the field together to discuss/define directions for research on the first and third phase of the affective loop, that is the design of system elements and responses that aim at fostering learning by eliciting favorable affective states. In particular, the workshop will try to address the following questions:

- When/why is it useful to include affect in an ILE? Can we begin to develop a taxonomy of learning contexts under which it is critical or useful to include affect in an ILE?
- What types of affective elements, and why, should be included in the affective loop (e.g. basic affect such as like/dislike reactions; low-level dimensions of emotional states such as valence and arousal; instantaneous emotions such as joy, eureka, anger; longer term emotions/moods such as frustration, boredom, flow)?
Can we identify a mapping between the taxonomy of affect-demanding context and the most appropriate level of affective capabilities that the ILE should have (i.e. responding to/expressing specific emotions, or moods, or just valence, etc.)?

Which are the designs and behaviors that can help an ILE foster learning through affect? Are the relevant theories/findings that we can use in psychology, cognitive science and education?

What are the best means of evaluating an affective ILE? What elements should we evaluate (e.g. learning, user satisfaction, motivation, engagement)? Are there innovative evaluation methods that should be considered to evaluate the potential and needs of affective ILE?

By addressing these issues in a mixed-mode, informal set of interactions, we hope to explore alternative methodologies for closing the affective loop in ILE, identify key problems to address, and contribute to advancing the state of the art of this exciting area of research.

July, 2009
Cristina Conati and Antonija Mitrovic
Program Committee

Co-Chair: Cristina Conati, University of British Columbia, Canada (conati@cs.ubc.ca)
Co-Chair: Antonija Mitrovic, University of Canterbury, New Zealand (tanja.mitrovic@canterbury.ac.nz)

Madeline Alsmeyer, University of Sussex, UK
Ryan Baker, Carnegie Mellon University, USA
Winslow Burleson, Arizona State University, USA
Rafael Calvo, University of Sydney, Australia
Sidney D'Mello, University of Memphis, USA
Lewis Johnson, Alelo Inc., USA
James C. Lester, North Carolina State University, USA
Rosemary Luckin, University of London, UK
Scott W. McQuiggan, North Carolina State University, USA
Roger Nkambou, UQAM, Canada
Helen Pain, University of Edinburgh, UK
Ana Paiva, Instituto Superior Tecnico, Portugal
Kaska Porayska-Pomsta, University of London, UK
Genaro Rebolledo-Mendez, University of Sussex, UK
Ma. Mercedes T. Rodrigo, Ateneo de Manila University, Philippines
# Table of Contents

Designing Affective Support to Foster Learning, Motivation and Attribution  
*Ivon Arroyo, Kasia Muldner, Winslow Burleson, Beverly Woolf and David Cooper*  
1

Two Approaches for the Design of Affective Computing Environments for Education  
*Rafael Calvo, Dorian Peters and Lawrence Peters*  
11

Constructing and Affective Path to Support Learners in a Distance Learning Environment  
*Makis Leontidis and Constantin Halatsis*  
20

Problem Finding in Academic Writing with Affective Tagging  
*Satoshi Suzuki, Aiko Shiraishi and Hiroaki Suzuki*  
30

Action Decomposition and Frustration Regulation in the Assisted Execution of Difficult Tasks  
*Alessandro Valitutti*  
40
Designing Affective Support to Foster Learning, Motivation and Attribution

Ivon ARROYO a, Kasia MULDNER b,1, Winslow BURLESON b, Beverly WOOLF b, and David COOPER b

a University of Massachusetts, Amherst, USA
b Arizona State University, Arizona, USA

Abstract. Our work involves the design, implementation and evaluation of Affective Learning Companions, real-time computational agents that infer students’ emotions and related states and leverage this knowledge to improve educational outcomes. In this paper, we focus on describing the affective support delivered by these learning companions. To date, unfortunately, there do not exist prescriptive theories on how to design the structure of affective support or how it should be tailored to suit individual student needs, even though these factors are critical in shaping learning and affect-related outcomes, making the design of affective support challenging. Our approach in addressing this challenge has been to involve human tutor experts in the affective support design process. Here, we present the wide range of affective scaffolding that we designed, which includes both verbal and non-verbal interventions, the stochastic algorithm for combining various affective interventions into complex messages aimed at fostering motivation, attribution and positive affect, and the evaluations we are currently conducting.

Keywords. affective support, learning companions, motivation, attribution

Introduction

Affect and related states such as motivation, empathy and attention play a fundamental role in influencing learning outcomes. For instance, learning fails to occur in the presence of negative affective states such as anxiety or anger [1], and is enhanced when empathy or support is present [2,3]; interpersonal relationships between teachers and students increase student motivation (e.g., [4]). Despite the critical role of affect in the educational process, however, to date the research community has focused on the cognitive dimension, at the price of neglecting affect, as is pointed by key researchers [5,6]. Consequently, there do not yet exist prescriptive theories on how to design computational affective support or how it should be tailored to suit individual student needs, even though these factors are critical in shaping learning and affect-related outcomes.

Our work involves the design and evaluation of Affective Learning Companions (ALC), real-time computational agents that infer students’ affective states and leverage this knowledge to increase student performance, affect and attitudes towards learning.

1Corresponding Author: Book Production Manager, IOS Press, Nieuwe Hemweg 6B, 1013 BG Amsterdam, The Netherlands; E-mail: bookproduction@iospress.nl.
Our key research question are: (1) how can a computational system determine, at any point in time, the affective state of a student, and (2) provide appropriate tailored support that will improve the student’s affective state and learning? In this paper, we focus on addressing the second question, and in particular, on the design of affective support delivered by ALC integrated into a computational learning environment.

One of the challenges with designing affective support is defining what this support should include, since as we point out above, there is a lack of fine-grained theories related to this issue. Research indicates that concepts related to affect such as self-efficacy play a critical role in influencing learning outcomes (e.g., [7,8,9]). In fact, emotional intelligence comprises self-motivation, empathy, self-awareness, impulse control and persistence [1]. Consequently, as is advocated in [5], we also argue that affective support should involve a wide range of scaffolding, spanning from empathetic responses, to attribution-based messages aimed at changing student feelings towards mathematics, to messages motivating students to shift their problem-solving strategy. To design this wide range of support, our approach has been to involve tutor experts in the feedback design process. We integrated the affective support into our test bed application, Wayang Outpost, a multi-media Intelligent Tutoring System (ITS) [10,11], thereby extending existing work by providing an ITS with a full spectrum of support aimed at fostering learning, positive affect, motivation and attribution.

We begin by describing some related work, and then present Wayang Outpost and its student model for assessing affect. We then describe the affective support we have integrated into Wayang, including the individual interventions and the stochastic algorithm for combining them into feedback messages. We conclude with an overview of the ongoing evaluations of this support.

1. Background and Related Work

Given that feedback plays a fundamental role in the educational process, the ITS community has been exploring how computational tutors can best provide it. To date, much of this work has focused on non-affect related factors (e.g., [12,13]). Recently, Shute [6] conducted a review of how various feedback-related attributes such as message complexity, timing and student expertise influence learning outcomes. As she points out, there is currently very little research on affective feedback and its role in shaping educational outcomes, and that there is a need to “systematically explore the affective components and outcome performance”.

A number of researchers have begun investigating these avenues. Some of this work is based on human-to-human tutoring sessions. For instance, Porayska-Pomsta et al. [14] analyze such sessions to determine how human tutors diagnose student affect and the actions that tutors take as a result. Likewise, Lehman et al. [15] analyze affective states students experience during learning, as well as tutor responses to students actions during one-on-one sessions. Boyer et al. [16] investigate the balance between providing motivational and cognitive (e.g., hints on the domain) feedback during tutoring. Their findings highlight that in some instances, cognitive and motivational feedback are at odds with one another: feedback incorporating both aspects increases self-efficacy but decreases performance. We should point out that while studies on human-to-human tutoring provide an important foundation for the design of affective support, translating the results
into computational models is complicated by the lack of fine-grained details, on for instance, wording or adaptation to concrete situations not described in the analysis, and/or understanding of how the findings translate to computer-to-human tutoring.

There is also some work on integrating affective support into ITSs. For instance, Zakharov et al. [17] describe an emotionally-intelligent agent that generates empathetic responses during problem solving, e.g., the agent appears sad when the student makes poor progress. The evaluation shows students in general prefer the affective agent over the non-affective counterpart. Burleson and Picard [18] describe a learning companion that mirrors students’ affective states through non-verbal gestures and provides affect-related verbal support, corresponding to informing students about Dweck’s “the mind is like a muscle” message [19]. The evaluation shows that tailored sensor-driven affective support influences students’ motivation and attitudes towards the tutoring application, and that this effect is mediated by gender.

2. Wayang Outpost

The test bed application for our research is Wayang Outpost, a multimedia ITS that trains students to solve challenging geometry problems of the type that commonly appear on standardized tests [10,11]. To answer problems in the Wayang interface, students choose the solution from a list of multiple choice options (typically four or five, see interface fragment in Fig. 1A). Wayang’s original form of support corresponded to the following two types. First, while students solved a problem, they could ask Wayang for hints. The hints contained information on geometry rules needed to solve the target problem; for a given problem, the hints would become progressively more specific until a bottom-out hint was reached. Second, between problems, i.e., after a student finished a problem and before she started a new one, Wayang presented progress reports that informed students about the accuracy of their problem solving (see Fig. 1B). The progress reports are designed to reify to students their productive and unproductive behaviours, and inform them about the consequences of their actions on their progress. We evaluated the utility of this approach [10], and found that compared to students who did not see progress reports, students who did see them learned better, had higher learning orientation, and attributed more human-like characteristics to the Wayang tutor.

We have been working on extending the range of feedback offered by Wayang to include a rich array of affect-based support. To deliver this new support, we decided to integrate learning companion agents into Wayang, who are designed to act like peers that care about a student’s progress, and offer support if he becomes frustrated or begins to lose interest. The underlying motivation for including peer-based companions instead of more traditional tutors is rooted in the extensive literature on the benefits of peer-to-peer tutoring (for an overview, see [20]). To design the companions, we collaborated with a cognitive psychology researcher, who provided us with concrete feedback on ways to make the characters more believable. Currently, we have two companions designed, Jane and Jake (Fig. 2), because we are interested in exploring how the gender of the companion influences the impact of affective support. Since we wanted the affective support to be tailored to students’ needs, our first step has been to develop a student model capable of assessing students’ affective states, that we now describe.
Wayang includes two student models: (1) a simple effort model that is used to assess the degree of effort a student invests to generate a problem solution, and is based on time per action (i.e., if a student invests very little time between actions, this implies guessing and hint-abuse); (2) an affect model that corresponds to a linear regression model that Wayang uses to predict a student’s current emotional state.

The affect student model, which we described in [11], is derived from data obtained through two evaluations we conducted in the fall of 2008 involving 38 high school students and 29 undergraduate students, respectively. Each evaluation involved students interacting with Wayang as part of their regular math class for 4-5 days. To obtain information on how students were actually feeling as they interacted with Wayang, we had Wayang prompt students on a set of four bipolar emotional axes, that define a total of 8 emotions. During the study, we also used a set of sensors to capture students’ physiological responses, including a pressure mouse, posture chair, skin conductance bracelet, and...
webcam supplemented with software for inferring affective states from facial features. All interface actions in Wayang and sensor data were logged.

To build the affect student model, we first identified a set of variables that we believed could be predictors of emotions, including (1) variables related to students’ interface actions, such as number of hints accessed (‘tutor context’ variables below); (2) variables related to the sensors, such as ‘sit forward’ events inferred from the posture chair (‘sensor variables’ below). We then relied on stepwise linear regression to verify variables that were good predictors of each type of emotion we captured, with a student’s self-reported emotion as the dependent variable, and the tutor context and sensor variables as the independent variables. We found that a combination of ‘tutor context’ and sensor variables resulted in the most accurate model for predicting student emotion (for details, see [11]).

3. Wayang’s Affective Interventions

We now present the affective interventions that Wayang’s learning companions deliver; ultimately, these will be tailored according to the student models described in Section 2.1. As we already mentioned, the design of affective support is challenging because there do not exist prescriptive theories on how to design such support or how it should be tailored to suit individual student needs. To address this challenge, our approach has been to involve teachers and scientists in the intervention design process, as well as to draw as much as possible from related work. In particular, the character messages were created in collaboration with several experts from the Center for Applied Special Technology (CAST) institute, which is a research and development organization that works to expand learning opportunities through universal design for learning (http://www.cast.org/). The experts included two research scientists both with doctorate degrees in education, who specialize in the role of emotions in the educational process in general, and addressing emotions during mathematics instruction in particular.

We met with the CAST specialists several times during the intervention design process, as follows (1) the CAST experts provided the initial intervention design and wording; (2) the first author then categorized these interventions according to type of intervention (e.g., empathetic, attribution, effort-affirmation), and also refined them (e.g., shortened them), relying on the related work as much as possible during this process (e.g., [21]); (3) the interventions were then sent back to the CAST team, who approved the changes, making revisions as necessary. We now present the set of affective interventions that were the product of this process.

3.1. Verbal and Non-Verbal Interventions Generated during Problem Solving

As we pointed out above, emotional intelligence involves factors such as self-motivation, empathy, self-awareness, impulse control and persistence [Goleman 1995]. Given that we want our learning companions to both appear emotionally intelligent and support the emotional intelligence of our learners, we have designed a wide range of affective support that incorporates empathy, attribution, strategy and effort affirmation. Wayang’s learning companions deliver this support while a student solves a problem. The support includes both verbal and non-verbal scaffolding.
### Table 1. Sample Affective Interventions

<table>
<thead>
<tr>
<th>Category</th>
<th>Type</th>
<th>Sample Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empathetic</td>
<td>Frustration or anxiety</td>
<td>“I often get discouraged when struggling with a math problem”</td>
</tr>
<tr>
<td>Empathetic</td>
<td>Anxiety</td>
<td>“You know - sometimes I get embarrassed when I get the answer wrong”</td>
</tr>
<tr>
<td>Empathetic</td>
<td>Frustration</td>
<td>“Don’t you sometimes feel frustrated with math problem solving? I do”</td>
</tr>
<tr>
<td>Attribution</td>
<td>General</td>
<td>“I found out that people have myths about math, like that only some people are good at math. The truth is that we can all be successful in math if we give it a try”</td>
</tr>
<tr>
<td>Attribution</td>
<td>Effort</td>
<td>“Keep in mind that when we are struggling with a new skill we are learning and becoming smarter!”</td>
</tr>
<tr>
<td>Attribution</td>
<td>No-Effort</td>
<td>“We will learn new skills only if we are persistent. If we are very stuck, let’s call the teacher, or ask for a hint from Wayang!”</td>
</tr>
<tr>
<td>Attribution</td>
<td>Incorrect</td>
<td>“When we realize we don’t know why that was not the right answer, it helps us understand better what we need to practice”</td>
</tr>
<tr>
<td>Effort Affirmation</td>
<td>Correct</td>
<td>“That was too easy for you. Let’s hope the next one is more challenging so that we can learn something”</td>
</tr>
<tr>
<td>Effort Affirmation</td>
<td>No-Effort</td>
<td>“Good job! See how taking your time to work through these questions can make you get the right answer?”</td>
</tr>
<tr>
<td>Strategic</td>
<td>Incorrect</td>
<td>“Are we using a correct strategy to solve this? What are the different steps we have to carry out to solve this one?”</td>
</tr>
<tr>
<td>Strategic</td>
<td>Correct</td>
<td>“We are making progress. Can you think of what we have learned in the last 5 problems?”</td>
</tr>
</tbody>
</table>

The non-verbal affective support corresponds to having Wayang’s learning companions mimic through their behaviors whatever the student is feeling, which is a form of an empathetic response. For instance, one non-verbal intervention corresponds to having the learning companion appear excited in response to student excitement (see Fig. 2, right). Behavioral mimicry is a form of empathy that is prevalent in social interactions (e.g., [22]), here we are exploiting that property to foster social relationships between the student and learning companion.

The verbal affective support corresponds to a variety of messages that are verbally expressed by Wayang’s learning companions. Note that these messages are designed to appear as if coming from a peer, instead of a tutor, who is solving the problems with the student in Wayang. We now describe the various types of verbal interventions that we have embedded in Wayang.

#### 3.1.1. Empathetic Interventions

The CAST experts suggested that it is particularly important to address students’ negative emotions during verbal empathetic responses, because such emotions interfere with learning. Consequently, the empathetic interventions are expressed by the learning companions to acknowledge and/or support students’ frustration and/or anxiety, with the hope of alleviating these negative feelings. If a student is frustrated and/or anxious, but it is not clear which, then the learning companion responds empathetically to address both emotions. If the student is only anxious, then the empathetic response is geared at addressing anxiety; likewise for frustration (see Table 1 for sample interventions).
3.1.2. Attribution Interventions

According to attribution theory, students don’t feel motivated to learn due to their attributions, i.e., beliefs, of why they succeed or fail at tasks (Weiner, 1986). Attribution training involves changing a student’s beliefs in the causes of his or her own failures and successes to promote future motivation for achievement, and can have a significant impact on learning attitudes [23]. We have the following types of attribution interventions (see Table 1 for examples):

- The general attribution interventions are generated to encourage students to change their attitudes and feelings about math and learning in general.
- The effort attribution interventions are generated when students are investing effort during problem solving but are struggling, and are designed to help students realize that this is a necessary by-product of learning.
- The no-effort attribution interventions are generated when students are not investing effort when problem solving, and are designed to help them realize that effort is necessary in order to learn.
- The incorrect attribution interventions are generated to motivate students after they get an incorrect response, by changing the way they think about errors.

3.1.3. Effort-Affirmation Interventions

The effort-affirmation interventions acknowledge when students are (or are not) investing effort during problem solving. In contrast to the effort-attribution interventions presented above, which aim to change students’ attitude towards effort during problem solving, the effort-affirmation interventions acknowledge effort or lack of with respect to a student’s problem solving entry. By having the learning companion affirm effort (or lack of), the hope is to both build a more realistic social bond between the companion and the student and to motivate the student. The effort-affirmation interventions include the following (see Table 1 for examples):

- The correct no-effort interventions are generated after a student invests no effort and gets the answer to a problem right, to acknowledge that the problem given to the student was too easy for them.
- The correct effort affirmation interventions are generated after a student invests effort and gets the answer to a problem right to acknowledge the student’s effort.

3.1.4. Strategic Interventions

The strategic interventions aim to motivate students to reflect on their problem-solving strategies, and to foster motivation for problem solving as a result, as follows (see Table 1 for examples):

- the incorrect strategic interventions are generated when students are not succeeding at getting correct answers during problem solving, and are designed to motivate students to change their general problem-solving strategy, i.e., think about why they are not succeeding.
- the correct strategic interventions are generated when students are succeeding at getting correct answers during problem solving, and are designed to encourage students to think about their problem solving strategy, i.e., to consider why they are succeeding.
4. Algorithms for Integrating Affective Interventions

Given the above set of various types of interventions, a question relates to how these should be presented to a student. While an option is to present each intervention individually, another is to combine the various types of interventions into an overall response. The advantage of doing the latter is that the integration of verbal and non-verbal feedback affords a more complex and potentially believable impression of the learning companion delivering the message. The disadvantage, however, is that the overall message length increases and there is a higher potential for repetition. We are experimenting with both approaches: (1) generating the interventions individually, according to a stochastic algorithm to increase variability, (2) combining the interventions into an overall message. The latter is accomplished by two algorithms we have designed, used (1) when a student’s answer is incorrect (see Fig. 3A); (2) when a student’s answer is correct (see Fig. 3B). The algorithms take into account student effort and/or affect to produce an overall message consisting of empathetic, attribution, strategic and effort-affirmation components, as appropriate. The algorithms include some stochastic properties, which are embedded to introduce variability in the learning companion’s affective behaviors and so increase its realism.

To illustrate the algorithm output, let’s suppose a student generates an incorrect answer really fast (indicating no effort); the student is frustrated. Intervention generated according to the above algorithm is as follows:

Figure 3. Wayang algorithms for integrating affective feedback presented in Section 3 into overall response: (A) algorithm used when student answer is incorrect; (B) algorithm used when student answer is correct.
Character looks frustrated for a few seconds. (Non-verbal Empathetic)
Don’t you sometimes feel frustrated with problem solving? I do. (Verbal Empathetic)
However, (Connector)
we will learn only if we are persistent. If we are very stuck, let’s call the teacher, or ask for help. (No-Effort Attribution)

Now the student again generates an incorrect answer, but the effort is not clear and the student is anxious. Intervention generated according to the above algorithm is:

You know ... sometimes I get embarrassed when I get the wrong answer... (Verbal Empathetic)
But (Connector)
I think it is important to have an open mind and the belief that one can do math. (General Attribution)

To refine the approach for generating interventions (e.g., individual messages vs. combined messages), we are currently piloting the various options.

5. Planned Evaluations and Future Work

Above we described the affective support we have designed and embedded in Wayang Outpost. This support includes both verbal and non-verbal interventions, which may be combined via stochastic algorithms we presented in Section 4 into complex messages aimed at fostering motivation, attribution and positive affect. Our next step is to evaluate the interventions. A general challenge is at what level to assess the affective support. One can take a broad approach and assess the impact of the interventions as a whole, without controlling for the impact of the individual components (e.g., visual vs. verbal feedback). An alternative is to conduct ablation studies, where we evaluate, for instance, the impact of each type of intervention. For the time being, we have chosen the former option, because we want to begin by exploring whether affective support has an impact in general before we tease apart the factors contributing to this impact (if any).

We are currently in the process of conducting a series of evaluations with Wayang, which include validation of its user models with new populations, and piloting of the interventions in preparation for the evaluation of the affective support, which will take place this spring with a series of high schools. In this study, all students will interact with Wayang over a series of sessions (typically 4-5) and all students will be asked to self-report on their emotions at fixed time intervals (the prompt to do so appears after a student solves a problem and before she starts a new one to minimize interruption). To evaluate the impact of Wayang’s affective support, our study includes three conditions: [1] students receive none of the above-described interventions (control condition); [2] students receive all of the above-described interventions but these are generated randomly instead of according to Wayang’s student models (random feedback condition); [3] students are provided interventions tailored based on their self-reported emotional state and/or effort invested. Our hypothesis is that the tailored affective support will be most effective in terms of learning and affect-related outcomes.
References

Two approaches for the design of affective computing environments for education

Rafael A. Calvo\textsuperscript{a}, Dorian Peters\textsuperscript{b}, Lawrence Peters\textsuperscript{c}

\textsuperscript{a}School of Electrical and Information Engineering, The University of Sydney
\textsuperscript{b}Faculty of Education, The University of Sydney
\textsuperscript{c}Department of Theatre and Dance, California State University, Fullerton

Abstract. This paper proposes two design approaches for developing affective tools. The first design is based on an information processing approach and describes a tool that utilizes classification algorithms for automated emotion recognition. The second design follows an interactionist philosophy and describes a visualization tool to support learner/instructor interpretation of affective feedback. These two approaches reflect two contrasting methodological arguments in HCI that have distinct impacts on pedagogy. We demonstrate these designs in the application area of actor training, due to the uniquely explicit link between emotion, emotion recognition and learning in this discipline. We describe a series of experiments, run with both professionals and non-actors, which demonstrate the feasibility of these tools for blended or intelligent learning environments.

Keywords. Physiological signal processing, data mining, affective computing, performing arts, acting.

1. Introduction

1.1 Designing Affective Learning Environments: Information processing and Interactionist approaches

Designing affective computing tools that support learning remains an open question. Herein, we propose, that drawing from the literature in Human Computer Interaction (HCI), we can develop design methodologies that are pedagogically driven and benefit from existing research. Specifically, we look at how two HCI-based design approaches \cite{1} can be applied to the design of affect-aware learning technologies. The information processing approach treats emotion as an entity similar to information, that is communicated from one person to another, and for which designers can build models used to improve user interfaces by increasing the accuracy of adaptation. In contrast, Boehner \cite{1} and others posit a second interactionist interpretation of emotions, in which they are co-created while subjects interpret visualizations of unspecified emotions.

The first approach has, by far, been the most widely employed up to this point. An example of researchers who follow this line of work are Gratch and Marsella \cite{2} who

\footnote{Corresponding Author: Rafael A. Calvo. School of Electrical and Information Engineering, Bldg J03, The University of Sydney, NSW 2006, Australia. E-mail: rafa@ee.usyd.edu.au}
have developed a computational model of emotion based on appraisal theory, that they use to make more 'believable' agents or virtual humans. EMA (Emotion and Adaptation) uses the appraisal factors that trigger emotions and coping strategies, and has been used to adapt virtual humans that interact with users through natural language in high stress social settings (war scenarios). Work by Conati [3, 4] and Boehner [5] offer further examples of this affect-as-information approach. The underlying assumption of many of these systems is that automatically recognizing the learner's emotional state can enhance a system's capability to interact with the user effectively. For Intelligent Learning Environment development, there is the additional pedagogically driven assumption that an 'emotionally intelligent' tutoring system (or any learning technology) can support students to achieve better learning outcomes. Managing emotions was not the intent of any of these systems.

The interactionist approach was described and compared to the affect-as-information approach by Boehner [1]. The first system – Miró [5] was designed following the information processing approach and used sensor data to feed a model that was then used to create a visualization of the 'mood' of a group of people. The second system – Affector, followed the interactionist approach and the idea that systems can appear intelligent and exhibit complex behaviour without complex representation and processing of information [6]. Affector is an adaptation of a video conference tool connecting two offices, each with a screen and a webcam. Sensors in each office are used to distort each image in ways that can be 'interpreted' as representing different emotions. What is different in this approach is that the rules that map sensor signals to specific distortions can be defined by rules created by the users themselves.

1.2 Actor Training as an area of application.

As Jeanne Moreau once explained, “Acting deals with very delicate emotions. It is not putting up a mask. Each time an actor acts he does not hide; he exposes himself.” An essential element of actor education involves helping students learn to elicit and embody genuine emotions and not to rely solely on appearances. In fact, when an actor relies exclusively on the appearance of emotion he is said to be “indicating” [7] which is viewed as false or stylised. Affect influences all types of learning, but it is uniquely in the performing arts that the importance of affect is also at the core of learning outcomes. In no other discipline is the link between emotion, emotion recognition and learning so explicit. Thus it stands to reason that dramatic arts education stands to both a) benefit from emotion recognition and affective technology and b) provide a unique opportunity for experimentation and insight to researchers in the area. Just how these technologies could be feasibly integrated into the classroom and what part they might play within learning activities is an expansive question, but initial findings from a preliminary series of experiments with two distinct approaches to affective computing design, and ideas for practical educational applications are discussed herein.

This paper explores the feasibility of a new type of educational technology specially designed for supporting performing artists, particularly actors. We are applying the two different approaches described earlier, using physiological sensor data as input to an affective computing system.

Section 2 of this paper reviews the pedagogical issues commonly raised in actor training that relate to the support offered by the tools. Section 3 describes a visualization produced with features extracted from physiological signals. These features represent a form of summary that, instead of relying on a computer-generated mapping, allows for student/instructor interpretation as part of an educational activity. Section 4 describes
recording and recognition algorithms as they are commonly used in the affective computing literature, discussing an evaluation of the accuracy of different classification techniques and of data from different subjects (trained actors and non-actors). This section also discusses how they could also be used in supporting the training of actors. In section 5 we conclude, discussing future work.

2. Pedagogical issues in actor training.

Previous research in Affective AIED has focused on supporting student learning with content specific matters. Emotion has only been another variable to be accounted for in complex student models, with the goal of improving their accuracy. In contrast, within the discipline of acting, emotion is at the core, so ‘emotional intelligence’ issues that remain obscured in other training contexts are exposed more openly in actor training.

Other related research has been around how to reliably collect information about students emotions [8] and the so called ‘emotional cartography’ or the visualization of intimate biometric data and emotional experiences using technology [9].

Much of modern actor training is based on the approach created by the renowned Russian acting teacher, Konstantine Stanislavski. Key elements of Stanislavski’s system, include relaxation, emotional memory and concentration of attention. Stanislavski describes a series of elements, but the tools discussed in this paper focus on the enhancements of these three.

The first element, relaxation, is critical to an actor’s success. “Even the slightest tension of muscles can paralyze an actor’s creative state.” [10] Actors are famous for their process of relaxation before going on stage, and virtually every acting text will offer methods for finding this state of relaxed muscle tension, or neutrality.

The second element, referred to as emotional memory, makes use of an actor’s real life experiences in the elicitation of on stage emotions. “The actor must be capable of bringing out the imprint of a past experience and of making it respond to the conditioned stimulus on stage at the moment he needs it.” [10]

Finally, the third element is concentration of attention, which is also referred to in Stanislavski’s writings as “public solitude.” According to Stanislavski, actors must develop the skills to, through concentration of attention, represent actions and emotions realistically, while being watched by any number of audience members. Moore states that, “The more an actor exercises his concentration, the sooner it will become automatic; finally, it will become second nature to him.” (p34)

The practice and development of relaxation and concentration of attention, and the power of these abilities in finding and clarifying emotions through sense memory, is an essential component of Stanislavski’s system. A tool that validates an actor’s work in this area, providing physiologically based feedback could prove effective in the often vague and inconclusive nature of this work. A deeper level of relaxation may be found and focus of attention may be strengthened, facilitating the discovery of emotional memory, which is often elusive and fragmentary. A tool that allows an actor to find consistent personal emotion may help young actors to develop this technique, and make emotional memory more easily accessible. Initial experiments in supporting these elements with affective computing tools are described below.
3. Visualizations for an Interactionist approach

Boehner described [1] design principles that could be used to design this type of systems:

- Affect as a social product. The examples show how the idea that emotions can be constructed in the process of a person interacting (with a human or a machine).
- Elicitation and Interpretative flexibility. The information processing approach leads to applications where a model is trained with data from subjects who must assess their emotions as one or a combination of a finite set, or as a point in a multidimensional space (e.g. valence / arousal). These models are then used in other instances, possibly by other people. The interactionist model allows users to create their own meanings, and change them in other circumstances.
- Supporting an extended range of communication acts. By not constraining users to a finite number of emotional expressions, they can create their own communication acts.
- Experience centered: instead of making systems more aware of emotions, make people more aware of emotions through system design.

Visualization tools that provide feedback have been developed to support teaching in a variety of scenarios, from group activities to writing support. Using the physiological sensor data as input data, we can produce 2 or 3 dimensional representations of physiological features that describe a 'state' that the actor and coach can use to discuss.

As a proof of concept, sensors where used to record heart activity (electrocardiogram - ECG), face muscle activity (electromyogram - EMG), and skin conductance. Picard [11] developed a recognition system using these features and blood volume pressure (photoplethysmyograph). We asked 3 subjects to enact 8 emotions (same used in [11]): (no emotion, anger, hate, grief, platonic love, romantic love, joy, and reverence). The recordings were made with all the benefits of a lab setting, where subjects where located in a quiet room, with relative privacy and were not required to do any simultaneous cognitive activity (i.e. remember the lines of a play). The subjects were aware that the project was about studying emotions. Naturally, as we approach more realistic scenarios these factors may all affect the performance of the technologies.

After subjects were prepared for the study, the eight emotions mentioned earlier were elicited in order. Each emotion was elicited for a three minute period, separated by a period of rest. In this study subjects were not told exactly what was meant by each emotion (beyond its name) allowing individual, subjective, interpretations of each emotional label. After each emotion was elicited, subjects were asked to rate the emotion in terms of Arousal, Valence and Dominance on the Self Assessment Manikin pictorial scale [12]. Three sessions were recorded for each of three subjects (an experienced male actor and a female and male with no acting experience). Each session took place on a different day. The sessions with Subject 1 were recorded at 40Hz, while the sessions of Subjects 2 and 3 were recorded at 1000Hz, after the decision was made to see the effect of a higher sampling rate on classification performance.

The signal data for each emotion was organised into 30 overlapping 30 second windows. For each window 120 features were extracted using the Augsburg Biosignal Toolbox [13]. The features extracted were primarily the mean, median, standard deviation, maxima and minima of several characteristics in each signal.

For the sake of visualization, the 120 features where reduced to two and plotted using Fisher Projections [12]. Figure 1 shows the clusters of eight emotions produced by one of
the subjects in the study discussed in Section 4. Other feature selection techniques such as Principal Component Analysis (PCA) may be used. In this representation there is no actual model mapping features to emotions, the subjects would know which emotions they intended to feel and would see the points appearing on the screen. If the student is not focused the clusters might be more spread, since the spread might be interpreted as not being able to ‘sustain’ an emotional state. In collaboration with the coach the meanings of overlaps between clusters would be discussed. It is in the process of discussing with the coach, and reflecting on what they see that they learn.

![Figure 1 Fisher projection of physiological representation of 8 emotional states.](image)

The tools proposed would not be developed with the goal of making a naïve attempt at classifying ‘good’ or ‘bad’ acting. This is neither realistic nor desirable. Instead, what we envision are tools that could provide visualizations of emotions that could be helpful in prompting student reflection on their own practice. For example, the tool might provide visualizations that reveal attempts at eliciting emotions to be distinct or ambiguous, embodied with greater strength or less conviction, etc. These could encourage student reflection and provide support for teacher observations and suggestions. For example, an instructor may observe that there is little notable difference between a student’s portrayal of anger, hatred and anxiety. She could discuss this with the student, but if she were able to demonstrate this lack of distinction by showing a visualization of this overlap, the problem is suddenly more clear. This could then prompt an exercise in which the student takes a closer look at their definition of these emotions. How does one differentiate hatred and anger? The teacher might suggest exercises to improve focus and concentration, or improvisations based on situations that elicit each discretely. It could also prompt a deeper look at the student’s techniques for eliciting these emotions. Is she relying on a personal memory or experience, for example? Perhaps another technique would be more effective in this situation or for this emotion. After making amendments to his approach, a new session may reveal altered results. Here again, teacher feedback could be sufficient to communicate the student’s progress, but a visualization that now shows three clearly differentiated emotions where once they all overlapped could provide
uniquely effective reinforcement. In short, the tools could provide a concrete and visual representation of concepts usually confined to the abstract and subjective world.

It is also worth mentioning that these tools could provide extensive opportunities, not only for supporting actor training, but more broadly for supporting research into acting as a dramatic discipline, for example by providing a new method for understanding how acting works internally and how acting emotion manifests itself physiologically; for the comparison of different acting styles and techniques (e.g. Stanislavsky v. Alba Emoting); how the success of different techniques varies from actor to actor and among different dramatic styles (e.g. American Realism v. German Expressionism) or for different media (e.g. theater v. cinema), and how this correlates to perceptions of good or bad acting. Of course, a better understanding of acting as a discipline will also lead to improvement in actor training in the long run.

4. Affect-as-information and Recognition using physiology

Once the student has participated in an emotion elicitation session or sessions, this data can be used to train a model used for recognition and creating tutoring functionalities. If the system is able to accurately recognize emotions as he participates in a scene, a play or other activity a whole realm of opportunities open up for analysis of emotional dynamics.

A possible way of using an emotion recognition in a actor training scenario, assuming that trained actors ‘feel’ (physiologically) an emotion in similar ways would work as follows: 1) a group of trained actors elicit a number of emotions over different sessions. 2) data collected is used as a training set (‘gold rule’) for an automatic recognition system 3) The student enacts some of these emotions in a particular scene and the system points to ‘misses’, where the wrong emotion has been recalled, possibly due to wrong emotional memory, or has not been sustained for the whole scene due to lack of concentration.

For a system created with this design approach accuracy of the recognition is crucial. Picard [11] developed one such recognition system, training machine learning algorithms with data from 20 sessions where 1 subject elicited the same 8 emotions listed earlier, achieving 81% classification accuracy. They analyzed a total of 40 features, identifying the best 11 features for their subject. Although the set of features may differ for other subjects, the methodology described is quite general.

We investigated some of the effects on classification results of variations in factors such as: number of sessions, number of subjects, and algorithms used for classification. Eight classification algorithms were evaluated using 10-fold cross validation in WEKA [14]:

1. ZeroR: predicts the majority class in the training data; used as a baseline.
2. OneR: uses the minimum-error attribute for prediction [15].
3. Function Trees (FT): classification trees that could have logistic regression functions at the inner nodes and/or leaves.
4. NaiveBayes: A standard probabilistic classifier using estimator classes. Numeric estimator precision values are chosen based on analysis of the training data [14].
5. Bayesian Network: using a hill climbing algorithm restricted by sequential order on the variables, and using Bayes as optimisation criteria.
6. Multilayer Perceptron (MLP): using one hidden layer with 64 hidden units.
7. Linear Logistic Regression (LLR) using boosting.
8. Support Vector Machines (SVM): Finds the maximum margin hyperplane between 2 classes. Weka’s SMO with polynomial kernel was used [16].
Table 1 shows the accuracy of the best classifier (SVM). The columns 1,2,3 represent the results for each single session, while a dataset with the 3 together is in ‘combined session’. The rows show results for each subject with data standarized at 40Hz sampling.

<table>
<thead>
<tr>
<th>Subject</th>
<th>1 – 40Hz</th>
<th>2 – 40Hz</th>
<th>3 – 40Hz</th>
<th>Combined Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – 40Hz</td>
<td>96.3%</td>
<td>92.1%</td>
<td>95.4%</td>
<td>80.4%</td>
</tr>
<tr>
<td>2 – 40Hz</td>
<td>94.2%</td>
<td>97.5%</td>
<td>95.8%</td>
<td>74.7%</td>
</tr>
<tr>
<td>3 – 40Hz</td>
<td>90.5%</td>
<td>95%</td>
<td>92.1%</td>
<td>68.1%</td>
</tr>
<tr>
<td>All Subjects (40Hz)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>42.2%</td>
</tr>
</tbody>
</table>

Table 1: Percentage of samples correctly classified for data sets using the SVM Algorithm.

An underlying hypothesis of this study was that the classifiers’ performance gives an indication of an internal ‘consistency’ of the data. If the performance is bad for all algorithms, the data is harder to model. A number of specific experimental issues arise, including:

1. **Intra-Subject, Single Session  – concentration**
   Classifiers trained and used on a single session for a single subject are the most accurate, and in that sense represent what we would like to achieve in a more general use recognition system. On the other hand we hypothesized when a classifier shows less accuracy for a particular subject/day it could be due not just to the classifier’s properties but also to the fact that that the data is less consistent. In pedagogical terms this corresponds to Stanislavski ‘concentration’, described earlier.
   The results (Table 1) showed that the physiological signals can be used to produce classifiers with over 90% accuracy. They did not show significant differences for the single session datasets of the professional actor than the single-session datasets of the novices. Although we cannot generalize from these results, they are interesting and require further evaluation.

2. **Intra-Subject, All Sessions – emotional memory**
   The consistency with which a subject ‘feels’ an emotion between different sessions is similar to Stanislavsky’s emotional memory. In general, each subject will elicit emotions in different ways on different days. To build a classifier and to test it on data from a single session means excluding the factors of intra-session variation. A subject specific classifier can be trained and tested by combining data from a number of sessions. By combining the data from the three sessions for each subject into a single data set, the classifiers’ accuracy indicates how inter-session variation in emotional elicitation reduces the accuracy. This is probably caused by differences in the appraisal of emotion, intensity and quality of the elicitation.
   The results in Table 1 show that the same classifier trained/tested on the multi-session data set is less accurate than the ones trained on a single session, but still achieves accuracy above 68%.
   We can see in Table 1 that the accuracy on the expert actor (subject 1) data, at 40Hz, is much higher than the equivalent for other subjects. Comparing the results of individual sessions, some emotions were consistently poorly classified, others consistently well classified, and others varied from session to session. Platonic love and romantic love stand out as emotions that are often misclassified, while anger and the no-emotion baseline were consistently well classified. Since the data collected for the 3 subjects is limited we do not
make general inferences, but indicate that these aspects of the confusion matrix of our very accurate classifiers, are worth further study.

The consistency with which an emotion is elicited appears in the data set combining all of a subject’s sessions. Subject 3 (a novice), for example, shows very high classification success in each session individually, but displays the lowest classification success in the combined data set. This suggests that for each individual session, the consistency of emotion elicited for each 3-minute block was very good (the ‘concentration’), but that the character of emotion elicited from session to session was not as consistent (the emotional memory). Subject 1 (the expert actor), in contrast, shows greater variation within individual sessions, but better consistency across the three sessions.

5. Conclusions

The potential for creating education support tools for actor training that would incorporate emotion recognition technologies and emotion visualisation was discussed. Two design strategies, the affect-as-information most commonly used in the literature, and the affect-as-interaction recently proposed in the HCI literature where used to design two types of tools. Although the ideas described here have not yet being used in actual teaching scenarios, the data collected from professional actors and teachers (co-author in this paper) and novice non-experienced actors provides the first glance at the type of acting training systems that could emerge from the affective computing research. The two systems might complement each other and the efficacy of each is something to be determined. The similarities and distinctions between two approaches would be applicable to other teaching domains.

For the interactionist approach a visualization tool that presents a 2 dimensional representation of the physiological signals produced while eliciting different emotions showed data clustered for each emotion. Actors would be expected to interpret this data, and it is in the reflection process where learning would be expected to occur.

For the second type of systems, automatic recognition is required. The possibility of using real time recognition was evaluated. The strong consistency of classifier success (> 90%) across the nine primary data sets reported (Table 1) adds to the evidence that accurate classifiers are possible. The high accuracy shows that distinct emotions produce distinct and therefore classifiable physiological signals, and that, perhaps unsurprisingly, self-elicited emotions are particularly distinct and classifiable for the expert actor for which significantly better accuracy was achieved.

A noteworthy result was the consistency of misclassification within a subject’s data sets. Subject 3’s romantic love samples were often misclassified as joy, and all subjects showed some misclassification among negative emotions; anger, hatred and grief. The accuracy of the classifiers is reduced when data for all 3 sessions is combined which highlights the effect of intra-day variation noted by Picard [2].

This study provided a first step, but we plan to have a more accurate evaluation with more subjects, of how key factors such as the acting experience, number of sessions, number of subjects, and algorithms used for classification, play a role in the accuracy of classification. While this study lays down an important foundation for recognising the importance of these factors, a complete understanding of the ways they affect the results can only be properly obtained through more detailed studies. Future studies will look
deeper at the difference in results for experts and novices and how visualisation tools could most effectively be used to support student actors.

Acknowledgements

The authors are grateful to Iain Brown for his help in the recording and processing of the data.

References

Constructing An Affective Path To Support Learners In A Distance Learning Environment

Makis Leontidis and Constantin Halatsis
Department of Informatics and Telecommunications, University of Athens
leon@di.uoa.gr, halatsis@di.uoa.gr

Abstract. This paper presents an ontology-based method for the selection of the suitable sequence of affective tactics in the context of MENTOR. MENTOR is a Web-based Adaptive Educational Environment where the learner’s needs and preferences are supported using an ontology-based approach during the learning process. Firstly, the proposed method uses an ontology in order to store in the Learner Affective Model (LAM) the learner’s Affective Knowledge. Then, based on this ontology an appropriate learning path through the educational material of the MENTOR is constructing. This path is called affective learning path because it is formatted by a set of suitable affective tactics. In this way we form a schema where the Affective Information can be represented efficiently and exploited properly in order to maintain the efforts of the learner.

Keywords., Affective Computing, Distance Learning, Pedagogical Issues.

Introduction

Web-based Adaptive Educational Systems (WBAES) in their majority develop their educational dimension, based only on cognitive parameters such as learning styles, without taking into consideration the emotional factors that are related to the student’s mood and personality. Many Web learning designers realize that this omission deprives the education from an important pedagogical dimension. Thus, they conceive the necessity to turn their attention in affective subjects which influence the learning [1],[3].

Recently, research in computer science has begun to take emotions into account, because their influence in perception, reasoning, decision-making and learning is considered catalytic. A new field which is called “Affective Computing” [10] and is located in the scientific area in the intersection of artificial intelligence, cognitive psychology and physiology, has come to surface with the promise to cover this deficiency and offers a wide range of methods, techniques and applications which take into account affectivity.

According to the above perspective, MENTOR which is a WBAES uses an Affective Module in order to recognize the affective state of the student during his interaction with the educational environment and thereafter to provide him with a suitable learning strategy constructing in this way an affective learning path [6]. The operation of MENTOR’s Affective Module is based on the FFM [4] and the OCC
model [9] providing the system with the essential “emotional” information in order to
determine the strategy of learning in collaboration with the cognitive information.

The remainder of the paper is organized as follows: In section 1 we first introduce
the basic concepts of our framework. Section 2 presents the MENTOR’s Affective
Module while Section 3 analyzes the ontological representation of the affective
information. Sections 4 and 5, respectively, provide the process of the affective
learning path construction and an illustrative scenario. We conclude in Section 6.

1. Basic Concepts and Motivation

1.1. Affective computing, Emotions, Mood and Personality

The term Affective Computing involves the intention of Artificial Intelligence
researchers to model and incorporate emotions in intelligent systems. It is a novel and
important topic for the field of human computer interaction in order to improve quality
of communication and transaction intelligence between human and computer. It is
Picard [10], who coined the term affective computing. She defines affective as the
“computing that relates to, arises from or deliberately influences emotions”. Based on
this definition, an affective system must be capable of recognizing emotions, respond
to them and react “emotionally”.

In the conceptual map of affective computing, emotions play a predominant role.
Emotion is analogous to a state of mind that is only momentary. Although many efforts
have been made, there is not an explicit definition for emotion. It is easy to feel, but it
is hard to describe it. There are still basic questions in the emotion theory such as why
do we have emotions, what exactly causes them, how could we control them effectively,
but satisfactory answers are forthcoming. According to Ortony, Clore and Collins [9],
emotions are valenced reactions to events, agents, or objects.

Another important concept in the terminology of affective computing is the term
mood. Mood is a prolonged state of mind, resulting from a cumulative effect of
emotions. Mood differs from the emotion because it has lower intensity and longer
duration. It can be consequently considered that mood is an emotional situation more
stable than emotions and more volatile than personality. Scherer [11] mentions that
mood is an affective state of low intensity but long duration, which is incurred without
evident reason and is formulated in relation to person’s subjective sensitivity.

In affective computing the particular occurrence of emotions and the consequent
expression of mood are assigned to some extent to the individual characteristics that
distinguish one human being from another. These characteristics determine the
personality of a person which is related to the person’s behavior and mental processes
and has a permanent character [11]. It would be considered that personality refers to the
determinant and predictable attributes and behaviors by which people are identified and
categorized. Emotions and moods are connected with the term of personality by the
name of traits or factors. For instance, optimist, imaginative, nervous, envious, rational,
are some personality traits which personify a person.

1.2. The OCC and the Five-Factor models

Despite the significant theories that have been proposed for affective computing, the
two major theories, where the majority of affective systems are relied on, are the
cognitive theory of emotions (OCC) which is related to the origination and the appraisal of emotions and the Five Factor Model (FFM) which is connected to the explanation and the prediction of a person’s behaviour according to his personality.

In order to explain the origins of emotions and to describe the cognitive processes that elicit them, Ortony, Clore and Collins [9] formulated the OCC model. Regardless of the various attempts that have been made in order to define and explain sufficiently the emotional processes, this theory keeps a distinctive position among them. According to this theory, in connection to a person’s perception of the world, his emotions can be elicited. This process is named appraisal and the OCC model assumes that the emotions can be triggered by the assessment of three perception aspects of the world. These aspects are events, objects and agents. That is, all emotions engage a kind of positive or negative reaction to the way the world is conceived. The intensity of the affective reactions determines whether or not they will be experienced as emotions. According to this point of view, the OCC model has been integrated in many affective computational systems with the aim of recognizing the user’s affective state and implementing emotions in machines.

The second significant theory that is used for the integration of affective systems is the Five Factor Model. This is the most known model of personality and results from the study of Costa and McCrae [4]. It is a descriptive model with five dimensions (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) and views the personality as the set of all those characteristics that distinguish one human being from another. Due to these dimensions the model is also called OCEAN model. The FFM provides us with a reliable way in order to connect a student’s personality with his mood and emotions that he possibly experiences during the learning process. This is very useful because we are able to initiate student’s emotional state and select the suitable pedagogical strategy.

1.3. Ontologies and Representation of Affective Knowledge

A modern approach in relation to the representation of the affective information is the use of ontologies. Ontological knowledge representation contributes to the building of consistent user models and obtains the basis for interoperability support to authors and system developers especially in the complex knowledge domains of the WBAES [5]. The use of ontologies to the representation of affective knowledge in educational systems has the benefit that allows the student model to handle explicitly information related to the student’s goals, his prior knowledge, his individual cognitive abilities, his emotional states and his interaction with the system [8]. In the usage of ontologies in WBAES significant attention must be paid to the design, maintenance, integration, sharing, re-using, and evaluation of the ontology.

An ontology can be defined as a specification of a conceptualization. It is a formal way to represent the specific knowledge of a domain, providing an explicit and extendable framework to describe it. Therefore, it comprises a technique for describing formally and explicitly the vocabulary of a domain in terms of concepts, classes, instances, relations, axioms, constraints and inference rules. Ontologies represent knowledge in taxonomies, where more specific concepts inherit the properties of those concepts which they specialize [12]. A number of different languages have been used to define an ontology. Among them the most important are the RDF, RDFS, and OWL which base their expressivity on the XML syntax. We exploit the advantages of the ontological representation in our model to set the vocabulary, properties, and
relationships for learning and pedagogical concepts under an affective perspective. Taking advantage of the above, we use an ontology of emotions and affective tactics in order to achieve a formal representation of the Affective Information in the LAM.

2. The Architecture of MENTOR’s Affective Module

The architecture of MENTOR’s Affective Module is presented in Figure 1. The Affective Module has three main components: The Emotional Component (EC), the Teacher Component (TC) and the Visualization Component (VC), which are respectively responsible for: a) the recognition of learner’s personality (PR), mood (MR) and emotions (ER) during the learning process, b) the selection of the suitable teaching (TG) and pedagogical strategy (PG) and c) the appropriate visualization of the educational environment. The combined function of these components “feeds” the educational system with the affective dimension optimizing the effectiveness of the learning process and enhancing the personalized teaching. The main purpose of MENTOR is to create the appropriate learning environment for the learner, taking into account particular affective factors in combination with the cognitive abilities of the learner offering in this way personalized learning [7].

![Figure 1. The basic architecture of MENTOR](image)

The EC provides the system with a dialogue that can elicit emotions depending upon the semantics and its context. This dialogue is used in every new session and defines the current student’s mood. Based on this dialogue the student’s mood is recognized either as positive or as negative. In our approach, good mood consists of emotions like joy, satisfaction, pride, hope, gratification and bad mood consists of emotions like distress, disappointment, shame, fear, reproach. So as to deal effectively with the emotions’ elicitation process, the EC uses the LAM where the affective information is stored. An affective ontology of emotions is used for the formal representation of emotions.

The architecture of the MENTOR is designed with equal respect to the cognitive and the emotional dimension of teaching as well. So, we consider that the TC which is in charge of the formation of teaching consists of two subcomponents, the Teaching
Generator (TG) and the Pedagogical Generator (PG) which are responsible for providing the cognitive and emotional tactic respectively. Combining the interaction of its two sub-components, the Affective Module forms the appropriate affective tactic for the student. Therefore, we use the term Affective Tactic (AT) so as to denote that the learning method which is suggested by the TC is a two-dimensional combination of cognitive and emotional guidance and support [7]. In this way, a traditional instructional tactic is enhanced with a motivational one.

3. The ontological representation of the Affective Information

The Semantic Web is based on the development of ontologies which enable the organization of the educational material around semantically objects of learning information. In this manner the knowledge representation can be enriched with a combined approach of information technology and educational pedagogy. This approach enables reasoning and inference logic which facilitates the modelling of knowledge and the selection of suitable teaching and pedagogical methods.

The above advantages of the ontological representation are exploited in our model in order to set the vocabulary, properties, and relationships for learning and pedagogical concepts under an affective perspective, the result of which can be a set of rich schemas. These schemas can be defined in a machine-readable way so as to be suitable for sharing among various systems, understandable between humans and machines and appropriate for meta-data and semantics encoding. In this way the educational objects of the MENTOR’s learning environment accumulate more meaning by the relationships they hold and the possible inferences that can be extracted by these relationships. As a result, an affective ontology of emotions and affective tactics is created in order to achieve a formal and proper representation of the LAM and the system’s learning strategies. Thus, we are capable of dealing efficiently with the affective factors which occur during the learning process.

The structure of the proposed ontology is in compliance with the OCC emotions classification [9] as well as the OCEAN model of personality [4] and has been adjusted suitably in order to attain the requiring domain knowledge and the pedagogical representation of our educational system. This ontology, which is an application-domain ontology and is called Affective Ontology, contains the necessary affective information to model and support specifically the MENTOR’s educational operations. Consequently, it has been built to recognize ten emotions which are: joy, satisfaction, pride, hope, gratification, distress, disappointment, shame, fear, reproach. The former five emotions compose the classification of positive emotions and are related to the positive student’s emotional state. The latter five emotions compose the classification of negative emotions and are related to the negative student’s emotional state. Because the construction of the ontology was based on the OCC model its concepts are defined in terms with this theory. For instance, the positive student’s emotional state is described as follows:

(POSITIVE-EMOTIONAL-STATE
 (SUBCLASSES
 (VALUE (JOY, SATISFACTION, PRIDE, HOPE, GRATIFICATION)))
 (IS-A (VALUE (EMOTIONAL-EVENT)))
 (DEFINITION (VALUE ("emotions or states, regarded as positive, such as joy, satisfaction, pride, hope, gratification"))))
In traditional learning we refer to teaching as denoting mainly the method which is followed by the teacher for the development of the student’s cognitive abilities. This definition implies also, however without stating it clearly, that the teacher is responsible for the emotional control and support of their students [13]. As it has already been stated, MENTOR takes this perspective into account in relation to the personality and the emotional state of the student, in order to provide him with a combination of cognitive and emotional guidance and support, which forms an Affective Tactic. In this sense, an AT is constituted of the Affective, Learning and Pedagogical Strategy, the Teaching Method and the Educational Material.

The structure of the proposed ontology is comprised of one main class the System_Affective class which is the root of the hierarchy. The System_Affective class represents the universe of MENTOR’s discourse that is everything which could be in system’s conceptualization and is divided into four sub-classes which are, the Learner_Affective_Model, the Learner_Personality, Learner_Affective_State and the Affective_Tactic sub-classes. The last sub-class represents the twenty ATs that have been already implemented in MENTOR. Every of these ATs is represented as a second layer sub-class, into these sub-classes. By performing this classification, the affective knowledge is represented appropriately and a corresponding affective knowledge base is constructed. For instance, the System_Affective class and the AT1 affective tactic sub-class are described as follows:

```
(SYSTEM-AFFECTIVE-CLASS
 (SUBCLASSES
 (VALUE (LEARNER-AFFECTIVE-MODEL, LEARNER-PERSONALITY, LEARNER-AFFECTIVE-STATE, AFFECTIVE-TACTIC)))
 (IS-A
 (VALUE (AFFECTIVE-INFORMATION)))
 (DEFINITION (VALUE ("information that define the individual characteristics of the learner and the suggested affective tactic")))

(AT1
 (IS-A
 (VALUE (AFFECTIVE-STRATEGY)))
 (DEFINITION
 (VALUE ("Congratulate the learner about a successful event "))) 
 (IS-A
 (VALUE (TEACHING-STRATEGY)))
 (DEFINITION
 (VALUE ("Problem Solving")))
 (IS-A
 (VALUE (LEARNING-STRATEGY)))
 (DEFINITION
 (VALUE ("Discovery learning")))
 (IS-A
 (VALUE (PEDAGOGICAL-STRATEGY)))
 (DEFINITION
 (VALUE ("Exercise")))
 (IS-A
 (VALUE (EDUCATIONAL MATERIAL)))
 (DEFINITION
 (VALUE ("Present narrative text and video"))))
```

We use the DL-OWL (Description Logic – Ontology Web Language) as a reasoning and inference mechanism to obtain the essential production rules, as well as
analyze the domain knowledge and interaction data. In this way, the formal and flexible representation of an AT can be achieved in relation to the learning goal of a student. The main advantage of this classification is that the inference process for the selection of the appropriate affective tactic can be performed effectively. The proposed ontology was implemented with the Protégé tool [2].

4. Constructing the Affective Path

With the term affective path is denoted the appropriate selection of the ATs’ sequence in order to preserve student’s mood upbeat. That is, taking into account the current affective state of the learner, a set of ATs is suggested to him in every step of his interaction with MENTOR, with the aim of changing and/or keeping positively his next affective state. Thus, the formatted affective learning path contributes to the adaptive regulation of MENTOR’s learning content according to the learner’s affective state, so that he can gain more knowledge and preserve positive mood from fitting contents.

MENTOR’s aim is to generate and suggest to the learner all the possible affective learning paths starting from his initiate knowledge state and guiding him suitably through the adequately selected topics of interest to the desired learning goal. The selection of such paths is performed in accordance with the affective information in the learner’s affective model (Figure 2). The consideration of the weighted routes contributes to the selection of the most appropriate from a short-distance point of view (Figure 3). For this reason we propose an algorithm which is based on the shortest path searching [14] in order to calculate the shortest affective learning path.

Firstly, the desired knowledge goal is obtained by the learner as a direct request while his current emotional state is recognized by the EC during his interaction with MENTOR. All this information is stored in the LAM, which is being updated during the learning process. Then the TC accesses the ATs’ Database and retrieves the suitable
Having found the objective AT, in the next step the TC generates all the possible affective paths moving backward up to a node which represents the learner’s current knowledge state. In case that such a node is not encountered, the TC searches backward along until an initiate knowledge state is found.

In Figure 3 are exemplified a portion of the affective paths and their precedence and succession relationships which are stored in the Affective Database. Given that the learner is familiar with the AT and the objective node that the student has to reach is the node AT, the process moves backward from the node AT along the relationships and stops when the AT is reached.

After the generation of the possible affective learning paths a Bayesian Network (BN) is constructed in order to optimize the already selected paths. This refinement is carried out by comparing the weighted arcs with the affective information of the LAM. Because the ATs are organized in the Affective Database as an AND/OR directed acyclic graph with the nodes connected by weighted arcs, it is easy and accurate for this information to be handled by the BN. The process of selecting the appropriate affective learning path is based on the algorithm which is described below:

```plaintext
for each node N of the AG do
    if arrow exists between i-j and i<>j then
        AT[i,j] = w_{ij};
    else
        AT[i,j] = null;
    end_if
    TEMP[i,j] = node_AG[index];
end_for
for each N do
    if w_{ij} > w_{i1} + w_{1j} then
        TEMP[i,j] = 1;
    else if w_{ij} = w_{i1} + w_{1j} then
        end_for
        for each w_{ij} do
            if node_AG[index] + AT[i,j] < node_AG[w_{ij}] then
                Begin
                node_AG[w_{ij}] = node_AG[index] + AT[i,j];
                Path[w_{ij}] := index-i;
                End;
            end_for
    end_for
for each i=1 to n
    node_AG[index] = i;
    minValue = node_AG[index];
    for each j from i+1 to n
        if TEMP[i,j] < minValue
            Index = j;
```
minValue = TEMP[i,j];
swap TEMP[i,j] and node_AG[minValue];
end_for

where AG is the directed acyclic graph of N nodes, AT is a matrix with the weight values \( w_{ij} \) of affective tactics between the nodes i and j. In case that i and j have no relationship a null value is assigned. The node AG[index] represents a path, where the node_AG[index] is equal to i if there is no connection between i and j, while Temp is a temporary matrix for the path indexes swapping. Taking into account this algorithm the candidate affective learning paths which are extracted according to the Figure 3 are:

- a:1,5,11,
- b:1-5-6-7-8-11,
- c:1-2-3-4-5-11,
- d:1-2-3-4-5-6-7-8-11,
- e:1-2-3-6-7-8-11

and their corresponding weights are:

\[ \sum W_a = 2.5, \sum W_b = 3.3, \sum W_c = 2.7, \sum W_d = 3.5, \sum W_e = 3.2. \]

Among these candidate paths, the most appropriate according to the shortest-path criterion is the path “a”, which is suggested to the learner as the proper pedagogical guidance.

5. An Illustrative Scenario

Let us examine, for example, the case of a student whose personality belongs to the Extraversion category, but his mood is recognised in the current session as negative. For this type of student the Teaching Generator has already selected an exploratory teaching method without examining his affective state. Before MENTOR applies this method, the Affective Module interacts with the Pedagogical Generator. The system takes upon making the student feel relaxed firstly by opening a short dialogue with him. Thereafter, it presents to him either a joke or a funny video clip (Figure 4), according to his preferences which are stored in the LAM. Finally, it motivates him either by encouraging him or by praising his abilities.

![Figure 4. The proposed Affective Tactic of a video clip’s presentation. The video clip presents in a funny way the problem of Cannibals and Missionaries](image-url)
6. Conclusion

In the present work we introduced an ontology-based method for the selection of the suitable affective tactics which essentially was developed in the context of MENTOR, a WBAES for personalized learning. The main purpose of MENTOR is to create and/or preserve a positive mood in the student, since this is a crucial factor for the learning process. The proposed method is anchored in Affective Computing models such as the OCC model of emotions and the OCEAN model of personality and supposed to enable accurate adaptive interventions according to the learner’s needs and preferences making use of a properly formatted affective learning path. The construction of the affective path is relied on a shortest path searching algorithm which creates candidate affective learning paths and selects the most efficient considering the least time and effort parameters in order to lead the student to his objective learning goal.

Furthermore, we are developing the MENTOR’s Affective Module bearing in mind to be independent from the specific domain model of educational systems, so that has the capability to be used by a wide range of them. Some preliminary experimental results are very encouraging for the further development of our proposed model. The Affective Module recognizes appropriately the affective states and needs of the learner so that he obtains the most suitable affective tactics. In advance research we plan to keep running an extensive experimental study conducting a web evaluation in order to testify more precisely the reliability of the proposed model. In this manner the accuracy of our system will be improved in order to be capable of supporting the learner with more suitable affective tactics.

References

Problem Finding in Academic Writing with Affective Tagging

Satoshi V. SUZUKI¹, Aiko SHIRAISHI² and Hiroaki SUZUKI³,a

¹Human Innovation Research Center, Aoyama Gakuin University, Japan
²Graduate School of Literature, Aoyama Gakuin University, Japan
³Department of Education, Aoyama Gakuin University, Japan

Abstract. Problem finding through reading is one of the crucial activities in academic writing. Unlike many previous studies that facilitated a learner's logical thought, we attempted to facilitate learner’s affective thought by developing an affective critical reading tool called EMU (Emotional and Motivational Underliner) because recent studies in cognitive science suggest that affective thought should facilitate appropriate problem finding. One of the characteristic functions of the EMU is affective tagging, and we predicted that affective tagging would prompt a learner writing an essay to find suitable problems from an article for her/his essay. Subsequently, an experiment was conducted to examine this hypothesis, and the results provided support for the affective tagging.

Keywords. Academic writing, critical reading, affective tagging, annotation

1. Introduction

Essay writing is one of the important academic activities especially for university students. Among many subprocesses of the essay writing, problem finding from texts is crucial for writing an original essay. Many researchers emphasise that critical thinking plays a significant role in problem finding [1]. However, it is important to realise two distinct processes in problem finding: problem recognition and problem formulation [13]. Problem recognition refers to a process where writers recognise a problem in an unarticulated way [13]. After this process, writers are expected to formulate the topic by critical thinking. According to the hypothesis of the two subprocesses in problem finding, critical thinking should not carried out at the first process, but at the second. For example, brain storming technique realises it by its famous principles: “focus on quantity (rather than quality)” and “without criticism” [12]. This implies that, at the beginning at least, unarticulated ideas might be suppressed by critical thinking.

In this paper, we first discuss the important role of emotions in the problem finding and suggest affective tagging for critical reading to find the problems. Next, we present a Web-based tool for critical reading that elicits readers’ affective responses to texts. We finally report the experimental results that show the effectiveness of the system in finding issues for essay writing.

¹Corresponding Author: Human Innovation Research Center, Aoyama Gakuin University, #151301 Goucher, 4-4-25 Shibuya Shibuya-ku Tokyo 150-8366 Japan; E-mail: ssv@hirc.aoyama.ac.jp.
The learners would be able to adequately argue about issues if they can have an interest in them. However, studies and practises of critical reading have excessively emphasised the importance of the rational thought in the problem finding. For example, Adler and Van Doren [1] pointed out the existence of two subprocesses in “analytical reading”: one is to recognise an author’s purpose and argument, and the other is to examine the consistency between them. According to their suggestion, the problem finding for essay writing should be thoroughly rational. However, some studies indicated that it would be inadequate or harmful for learners to logically argue issues only by logically understanding the claim of the author in the article [3,9]. These studies imply that the problem finding in a thoroughly rational way would fail to decide the problems appropriate for the learners. Thus, we focused on the role of affective thought in the problem finding, and attempted to have learners find the problems by an affective tagging strategy.

In order to engage in problem finding, it is important for learners to become aware of rudimentary and unarticulated ideas which tend to be ignored when the learners find the problems in a rational way. In studies of decision-making, Dijksterhuis, Bos, Nordgren, and van Baaren [5] pointed out that decision-making without deliberation resulted in better choice as compared to that with deliberation within complicated settings of a problem. Problem recognition without deliberation can help learners determine what strongly arouses their curiosity, what they cannot fully understand, and what they strongly oppose in the texts. To enable learners to become aware of such ideas, we emphasise the role of affect. Recent advances in cognitive science and cognitive neuroscience have revealed that emotion and reason reinforce one another. For example, the somatic marker hypothesis [4] explains that we can discard considerable irrelevant information and alternatives by emotion. By virtue of such an affective selection, the remaining information can then be considered by reason. Thagard [14] claimed that there should be a strong association between the process of scientific discovery and affective cognition (curiosity, happiness, anger, fear, etc.) based on a narrative analysis of scientific discoveries. We attempt to promote learners’ performance in problem finding by introducing triggers of affective thought with the tagging and annotation function in a Web-based system.

We utilise emerging Web technologies [2] to enhance learning for essay writing. In fact, some studies focus on examining the function of the Web-based software in order to collaboratively review drafts among learners [10,11]. We attempt to facilitate emerging Web technologies like tagging and annotation on a Web document for problem finding. This is because it is indispensable to leave annotations such as underlines and comments in margins on a text while critically reading it, as Adler and Van Doren [1] recommended. Tsubakimoto et al. [15] developed an annotation and mapping system called eJournalPlus and examined the effect of mapping on critical reading through an experiment. In this experiment, they determined the rule of annotation in a rational manner for all the participants. Unlike the approach of their study, we enabled learners to add the tags and annotations in an affective manner.

We expect users to accomplish problem finding for essay writing by reading an article. When a user leaves an annotation for an arbitrary string on articles registered on the EMU system, (s)he must add an affective tag to the string from a tag set provided by the system. The tag set consists of the following five tags with the accompanying usage descriptions (line colour and decoration in parentheses):

**Crucial.** Add this when you want to say “This part is crucial” (black).

**Really!** Add this when you want to say “Really, that’s right,” or “Good!” (blue).
Aha! Add this when you want to say “I haven’t realised this,” or “It’s great!” (green).
Yuk! Add this when you want to say “Hey, that’s wrong,” or “It’s awful!” (red).
Um... Add this when you want to say “You bet?” or “I don’t get it” (black with “?”).

These tags and description sets contain casual expressions in order to induce the user to respond to the article in an affective manner. We call the addition of one of these tags to the string, affective tagging. The user can add a comment to the string, and can also underline a string without providing comments.

The EMU system was implemented as a content management system using Ruby on Rails 2.0.2. The annotation interface, as shown in Figure 1, was implemented with JavaScript so that the user could use the EMU system only with a Web browser.

2. Method

We conducted an experiment to examine the influence of affective tags on problem finding in an article by the learner.

2.1. Participants

Fifty four Japanese undergraduate (junior and senior grade) students participated in the experiment. All the experimental instructions and materials provided were in Japanese. The experiment was conducted during a course on educational studies. The participants were randomly assigned to one of the three conditions explained below.

2.2. Design

The experiment had one three-level independent variable (between-participant variable). The following is a description of each level of the independent variable:

One-colour group (1C; n = 18) The participants in this group were told to underline and leave a comment in an article provided by the experimenter with one colour as they would usually underline and annotate paper handouts and books. This group was prepared as a control group for comparing participants’ annotation activity with the other groups. It should be difficult to attempt valid comparison between activity of a group of participants without any annotation and a group in which participants annotate on an article.
Five-colour (rational) group (5CR; n = 18) The participants in this group were told to underline and leave a comment in an article using the following tag set, instead of the one explained above:

**Important** Add this when *you think that this part is important* (substitute for Crucial).
**I agree** Add this when *you think that you can agree with this part* (substitute for Really!).
**I did not know** Add this when *you think that you are unaware about this part* (substitute for Aha!).
**I disagree** Add this when *you think that you disagree with this part* (substitute for Yuk!).
**I cannot understand** Add this when *you think that you cannot understand this part well* (substitute for Um...).

We prepared these tag sets and instruction of tag usage so that their meanings are equivalent to those of the original tag set and also to induce the user to rationally analyse an article. In contrast to affective tagging, we call the addition of these tags to the string, *rational tagging*.

Five-colour (affective) group (5CA; n = 18) The participants in this group were told to underline and leave a comment for an article using the affective tagging strategy. Taking into account the argument of affective critical reading and annotation, People do not make rational evaluation of the given information at the first time when they read texts. Rather, they judge it by intuition, which lacks the extensive analysis. If learners are required to use rational tags at this point, their intuitive ideas would be suppressed. This is because the rational tags requires rational justification. In contrast, when they use affective tags, they naturally express being-developed ideas, since these tags do not demand extensive analysis.

### 2.3. Procedure

Each participant was told to write an essay about the given article. First, (s)he tried to get accustomed to the interface of the EMU system by working on a trial document. Second, (s)he read an article that supported the increase in the gap between the rich and the poor [7]. Although the original article pointed out some flaws in the argument, the passages from which these flaws were pointed out were omitted during the experiment in order to observe how the participants could identify such flaws. Next, the participants had to add underlines and provide comments to the article based on the tagging strategy assigned to each experimental group. After performing these tasks, the participants answered a reading comprehension test without reading the article. Then, after performing these tasks the participants wrote the essay about the article. Finally, the participant was debriefed (the original article of Harlan [7] was presented) and thanked. The procedure was the same among these groups with the exception of the tagging strategies.

### 2.4. Measures

We adopted the following measures observed during the experiment:

**Number of underlines** We counted the number of underlines made by each participant.

**Length of comments** We counted the number of characters in the comments added to each underline.
**Tone of comments** Regardless of the added tags in the 5CR group and the 5CA group, two judges classified the tone of the comments added each underline to the following categories (correspondence rate was .97):

**Pro** The comments in the article towards which the participants showed an affirmative attitude or a surprising response were placed in this category (for example, “I think this is correct,” and “It’s true that ...”).

**Neutral** The comments in the article towards which the participants showed neither affirmative nor negative attitudes (for example, logical organisation, paraphrase and summary of, supplementary for, and irrelevant response to the underlined string) were placed in this category (for example, “Topic sentence”, “Suggestion”, and “I’ve come across it.”).

**Con** The comments in the article towards which the participants showed a negative or sceptical attitude were placed in this category (for example, “It isn’t exactly true,” and “It doesn’t stand to reason that ...”).

We averaged the number of comments for each category on the basis of an assessment of the measures since some comments were too ambiguous to interpret with the categories.

**Number of correct answers in the reading comprehension test** We counted the number of correct answers in the reading comprehension test. The test consisted of nine multiple-choice questions. Three of the questions were sentence-level, that is, questions that could be answered only by referring to a certain sentence. Another three of the questions were paragraph-level questions, that is, questions that required an understanding of the topic sentence in a certain paragraph when answering the questions. The remaining three of the questions were article-level questions, that is, questions that required an understanding of the thesis of the article when answering the questions.

**Tone of the essay** Two judges classified the tone of the essay in the following categories (correspondence rate was 1.00):

**PRO** The essays in the article towards which the participant showed an affirmative attitude were placed in this category.

**CON** The essays in the article towards which the participant showed negative attitude were placed in this category.

The judges indicated of which option the participant chose by circling the sentence instruction and/or the sentence in the essay, for example, “I (dis)agree with the author, because...” If these indications were not found in the essay, the judges first identified a topic in the essay and distinguished a thesis from the concession phrases.

**Evaluation scores of the essay** Two of the authors independently and heuristically evaluated the essay written by the participants using a 5-point scale (from 1: poor to 5: excellent). The evaluation was based on the following criteria: whether the participant tried to analyse the issues in the article and whether (s)he tried to argue the issues using the facts not mentioned in the article. We considered the average score between the two authors as the evaluation score. Pearson’s product-moment correlation indicated a high correlation between the two scores of the authors ($r = .85, p < .001$).
Table 1. Number of underlines with or without comments

<table>
<thead>
<tr>
<th></th>
<th>Without comments</th>
<th>With comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1C group</td>
<td>128 (60%)**</td>
<td>85 (40%)</td>
</tr>
<tr>
<td>5CR group</td>
<td>95 (40%)</td>
<td>140 (60%)*</td>
</tr>
<tr>
<td>5CA group</td>
<td>81 (40%)</td>
<td>121 (60%)*</td>
</tr>
</tbody>
</table>

*: p < .05, **: p < .001

Table 2. The number of the essays categorised by each tone

<table>
<thead>
<tr>
<th></th>
<th>PRO</th>
<th>CON</th>
</tr>
</thead>
<tbody>
<tr>
<td>1C group</td>
<td>4 (22%)</td>
<td>14 (78%)</td>
</tr>
<tr>
<td>5CR group</td>
<td>9 (50%)</td>
<td>9 (50%)</td>
</tr>
<tr>
<td>5CA group</td>
<td>3 (17%)</td>
<td>15 (83%)</td>
</tr>
</tbody>
</table>

Table 3. Number of underlines for each category of the tone of the comment

<table>
<thead>
<tr>
<th></th>
<th>Pro</th>
<th>Neutral</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td>1C group</td>
<td>7 (8%)*</td>
<td>45 (53%)**</td>
<td>33 (39%)</td>
</tr>
<tr>
<td>5CR group</td>
<td>50 (36%)*</td>
<td>32 (23%)**</td>
<td>58 (41%)</td>
</tr>
<tr>
<td>5CA group</td>
<td>27 (22%)</td>
<td>29 (24%)*</td>
<td>65 (54%)*</td>
</tr>
</tbody>
</table>

↑: higher, ↓: lower

*: p < .05, **: p < .01, ***: p < .001

3. Results

Annotation data in the article Table 1 indicates the number of underlines with or without comments. This result shows that participants in the 1C group added underlines to the article without comments more than those in the other groups. The chi-square test indicated there were significant differences among the cells ($\chi^2(2) = 22.60, p < .001$). The residual analysis on the cells of the number of the underlines without comments
indicated that the participants in the 1C group added the underlines without comments significantly more than those with comments \((z = 4.75, p < .001)\), while those in the 5CR group \((z = -2.44, p < .05)\) and the 5CA group \((z = -2.29, p < .05)\) added the underlines without comments significantly less than those with comments.

Table 3 shows the number of underlines for each category of the tone of the comment: Pro, Neutral, and Con. The participants in the 1C group tended to add Neutral comments; those in the 5CR group, Pro and Con comments; and those in the 5CA group, Con comments. The chi-square test indicated that there were significant differences among the cells \((\chi^2(4) = 42.92, p < .001)\). The residual analysis on the cells of the number showed that the cells in Table 3 with the italic numbers were significantly smaller than their counterparts in the other groups and those with the bold numbers were significantly larger than their counterparts in the other groups. For the 1C group, Pro comments were significantly less \((z = -3.97, p < .001)\) and Neutral comments were significantly more \((z = 5.14, p < .001)\) than the other groups. For the 5CR group, Pro comments were significantly more \((z = 4.09, p < .001)\) and Neutral comments were significantly less \((z = -2.58, p < .01)\) than the other groups. For the 5CA group, Neutral comments were significantly less \((z = -1.97, p < .05)\) and Con comments were significantly more \((z = 2.37, p < .05)\) than the other groups.

Figure 2 shows the mean and SD (the SD values are in parentheses) of the annotations data in the 5CR and 5CA groups categorised by tags. By comparing the data between the two groups for each tag, the participants in the 5CR group tended to use the Important tag and the I agree tag more than their counterparts in the 5CA group. In contrast, the participants in the 5CA group tended to use the Aha! tag and the Um... tag more than their counterparts in the 5CR group. We examined the difference in the tendency of tag use between the 5CR group and the 5CA group by a two-way ANOVA (split-plot design: the independent variables were experimental groups as the between-participant variable and tags as the within-participant variable) for the number of underlines and the length of comments. The analysis indicated that the interaction between groups and tags \(F(4, 136) = 4.46, p < .01\) was significant when considering the mean of the number of underlines. Then, we conducted a simple main effect analysis for each tag and the analysis showed a significant simple main effect of Important/ Crucial \((p < .01)\), I agree/Really! \((p < .001)\), and I cannot understand/Um... \((p < .05)\), but there was no significant simple main effect of I did not know/Aha! and I disagree/Yuk!. In addition,
there was significant interaction between groups and tags ($F(4, 136) = 4.24, p < .01$)
when considering the mean of the length of comments. The simple main effect analysis
for each tag was conducted, and the results showed a significant simple main effect of I
agree/Really! ($p < .05$) and I cannot understand/Um... ($p < .001$), although there was
no significant main effect of Important/Crucial, I did not know/Aha! and I disagree/Yuk!.

**Result of the reading comprehension test**  Figure 3 shows the number of correct answers
in the reading comprehension test for each level. The 1C group obtained a larger mean
of the total number of the correct answers ($M = 6.72, SD = 1.64$) than those obtained
in the other groups (the 5CA group: $M = 6.06, SD = 1.21$; the 5CR group: $M =
5.61, SD = 1.54$). We conducted a two-way ANOVA (split-plot design: the independent
variables were experimental groups as the between-participant variable and levels as the
within-participant variable) for the number of correct answers, but both the main effect
of the groups ($F(2, 51) = 2.59, n.s.$) and interaction between the groups and the levels
was observed ($F(4, 102) = 0.73, n.s.$) were not significant.

**Evaluation of the essay**  Table 2 shows the numbers of the essays categorised by each
tone. This indicates that the participants in the 5CR group who wrote PRO essays were
more than those who wrote PRO essays in the other groups. However, the chi-square test
indicated that there were no significant differences among the cells ($\chi^2(2) = 5.51, n.s.$).

Table 4 shows the number of comments categorised for each tone of the essay. While
the participants who wrote PRO essays added more Pro and Neutral comments than Con
comments, those who wrote the CON essays added more Con comments than Pro and
Neutral comments. The chi-square test indicated that there were significant differences
among the cells ($\chi^2(2) = 35.56, p < .001$). The residual analysis in the cells of the
numbers indicated that the participants who wrote PRO essays added Pro ($z = 3.06,
p < .05$) and Neutral ($z = 4.63, p < .001$) comments significantly more than Con
comments, while those who wrote CON essays added the Con comments significantly
more than the Pro and Neutral comments ($z = 6.92, p < .001$).

Figure 4 shows the mean and SD of the evaluation scores of the essay. The partici-
pants in the 5CA group obtained the highest essay scores, followed by the participants in
the 1C group. The participants in the 5CR group obtained the lowest scores. A one-way
ANOVA indicated that the main effect of the group was significant ($F(2, 51) = 5.77,
p < .01$), and multiple comparisons using Holm’s method indicated that there was a
significant difference in the essay scores between the 5CA and 5CR groups ($p < .01$).
No significant differences in essay scores were observed between the 5CA and 1C group
or between the 1C and 5CR groups.

**4. Discussion**

The percentage of underlines without comments by the participants in the 1C group was
higher than that by the participants in the other groups. They added more comments that
paraphrased, summarised, and supplemented the underlined string than the participants
in the other groups. The results imply that the tag sets in the 5CR and 5CA groups act
as a scaffold for generating comments. The participants in the 1C group should focus on
understanding the article rather than reading the article critically. Thus, the tag sets could
provide the participants with clues to generate comments on each underline.
The difference between the rational tagging and the affective tagging should appear during the process of adding comments. While the participants in the 5CR group added the underlines and comments with the *Important* and *I agree* tags, those in the 5CA group added them with the *Um...* tag. The *Um...* tag frequently used by the participants in the 5CA group would be relevant to emotions such as interest, curiosity, wonder, and avoidance of boredom that play an important role in problem finding [14]. Therefore, affective tagging should elicit problem finding through reading and essay writing.

On the contrary, despite the difference in tagging, there was no significant difference among the experimental groups in the results of the reading comprehension test. The results imply that tagging strategies used in this experiment should have little influence on a learner’s ability to understand the content of the text. In other words, affective tagging would permit the learner from accurately understanding the text.

The difference between the rational tagging and the affective tagging would appear in the process of essay writing. The relation between the tone of the comments and the tone of the essays was observed in Table 4 while the difference of the tagging strategies did not strongly influence the participant’s attitude towards the article as shown in Table 2. This suggests that the learner’s attitude towards the article while adding annotations should determine the attitude in her/his essay. On the other hand, the participants in the 5CA group tended to write the essay by finding issues from the article and utilising their knowledge in essay writing to argue these issues with a sceptical attitude. Thus, the response of a learner elicited by the affective tagging would prompt the learner to find issues from the article and connecting these issues to her/his knowledge.

Nevertheless, there are certain shortcomings with the analyses in this paper. First, we attempted to examine the influence of differences in tagging strategies on the learners’ problem finding with only one article. Different kinds of articles should be read based on these tagging strategies. Dijksterhuis et al. [5] also pointed out that decision-making with deliberation should be more effective than that without deliberation in the simple setting of problems. Therefore, articles in which rational tagging is effective may exist. Next, we could not distinguish the influence of tagging strategies from the influence of instruction for reading the article. Even without tagging, learners with affective reading strategy may be able to find better problems than the learners with rational reading strategy. Thus, the influence of reading strategy and tagging strategy should be independently investigated. Besides, the logical structures of the essay were not discussed. The evaluation of the essays based on the logical organisation differs from the evaluation focusing on the validity of the content [8], as was conducted in this study. For example, claims of the essay with concession should be more persuasive than those without concessions. Thus, the influence of the tagging strategies on the logical structure of the essay should be investigated.

Moreover, utilising the annotations data on the article for system response should be investigated. For example, an automatic suggestion function for problem finding can be developed based on the annotation data. The experimental results showed that the difference in tagging strategies elicited differences in learner’s attitude towards the article. It should be difficult for learners who tend to use tags and add comments that express an affirmative attitude towards the article (e.g., *Important*/Crucial and *I agree*/Really!) to have sceptical viewpoints in the article. Although there is a need to inspect individual differences among learners who use affective tagging should be needed, developing a system response to improve learning strategies can be a prospective research topic.
5. Conclusion

We focused on affective cognition in problem finding from an article on academic writing to elicit a learner’s problem recognition. We developed a Web-based critical reading system by adding affective tags and annotations on an article. We examined how the affective tags elicited the learner to find issues appropriate to her/his essay. The results suggested that the learners could find issues from her/his viewpoint and validly discuss the issues in the essay on the basis of the affective tags. The study on the affective cognition of problem finding can discover process of logical thought through affective cognition, as well as clues for developing a novel learning environment for essay writing.

Acknowledgements

This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Scientific Research (B), 20300271, 2008.

References

Action Decomposition and Frustration Regulation in the Assisted Execution of Difficult Tasks

Alessandro VALITUTTI 1,
Department of Cognitive Science and Education, Trento, Italy

Abstract. Task difficulty and frustration are two correlated aspects of task execution, particularly in the learning context. The connection between affective recognition and action decomposition is proposed as a way to reduce frustration occurring in the execution of difficult tasks. The proposed framework allows for an analysis of frustration regulation, which can be considered as a specific type of affective loop.

Keywords. Frustration, difficulty, action decomposition, assisted task execution

Introduction

In every complex or new activity there may be moments of impasse in which our repertoire of skills and tools turn out inadequate. Generally, when it happens, people reorganize themselves through the development of new strategies or tools, making themselves stronger or more prepared to face the problem. This holds particularly true for the learning activity [8]. For example, the preparation of a university exam requires the coordination of different tasks of study, writing, or exercising tasks, and difficulty occasions may be numerous. Nevertheless, the difficulty of the task induce stress and negative emotions such as frustration, an this affective state can be itself a source of feedback that increases difficulty. In other words, frustration may be a way to perceive the objective difficulty of tasks. If sufficiently intense or drawn-out, it has effect on attention and motivation then reducing the overall capability to perform the task.

In the present paper a specific way to address this issue will be discussed. In particular, the connection between affective recognition and action decomposition will be proposed as a way to reduce frustration occurring in the execution of difficult tasks.

Even if the relationship between difficulty and emotional state is more complex and not limited to frustration, in this work we want to emphasize the distinction and correlation between objective and subjective aspects of difficulty. Furthermore, we focus on frustration in order to individuate possible advancements on previous studies on frustration and human-computer interaction such as [5] and [4].

1E-mail: a.valitutti@unitn.it
1. Tasks and Executor Systems

In order to provide the conceptual background on which the main ideas of this study are based, some preliminary definitions will be introduced:

1. Task execution requires the existence of an executor system capable to execute (at least) one set of primitive actions. This term is employed in order to consider both human and machine executors of algorithms. This choice is a way to take advantage of theoretical results and ideas from computer science and artificial intelligence, in particular planning. The analogy between humans and machines as executor of algorithms allows us to focus on characteristics of human executor that make them different from machines. Human beings need to have not only knowledge and skills, but also motivation. In particular, emotional states strongly affect the capability and performance in task execution.

2. A task is simple, if it is represented by a primitive action and can be directly executed by the system, or complex, if it is completed through the execution of a total ordered plan (i.e., a deterministic sequence of primitive actions).

3. Given a repertory of primitive actions, there are more ways to execute a particular task. It can be represented as a set of total ordered plans or, equivalently, a partial ordered plan.

4. A set of task can be organized in an action hierarchy, in which each action is connected to the set of sub actions that allows for its execution. Therefore, a task can be represented as a structure of actions with a set of ordering constraints, in order to individuate the partial ordered plan. This representation have been employed, among others, in hierarchical decomposition partial order planners (HD-POP) [7], hierarchical task network planners (HTN) [2], and partial order hierarchical reinforcement learning systems [3].

5. The decomposition structure of a given task in term of primitive actions is hierarchical tree in which the root node corresponds to the task action and the leaf nodes correspond to the primitive actions.

6. Then the difficulty of a task, for a given executor system characterized by a set of primitive actions, can be defined as the “distance” between them and the main action, measured as the sum of lengths of all paths from the root to the leafs. The greater the distance is the more complex is the plan to perform the task. The access to strategic knowledge is crucial for the executive skill. It may have a motivational effect as well. In fact, without a plan whose steps are perceived as executable, people may believe they are not able to perform the task and the activity does not even start.

Some executor systems may have different sets of primitive actions. In the case of machines, the primitive action set depends on the structural level taken in account for the communication of instructions. For example, the interaction with a computer can proceed at the level of the machine code, the operative system, or applicative software. Nevertheless, even if a specific instructional context is fixed, the action set can change over the time. For instance, from the point of view of the user, a programmable computer can acquire the capability to execute new tasks through the installation of new software. It is interesting to observe that, for a system that can change the primitive actions, there are two degrees of freedom in the representation of a given task, respectively corresponding to the set of primitive actions and the set of total ordered plans for a fixed action set.
In the case of human executors, the ability to execute primitive actions can be defined as the attitude to perform a task automatically, without the need to pay attention to an explicit plan. Unlike machines, the set of primitive actions is extremely variable and correlated to several factors among which are individual skills, learning, and mental state. The analysis of next sections is focused on the dependence of primitive actions on the affective state of the human executor.

2. Assistance to Task Execution

A computational tool can be employed to assist the user in the execution of a complex task. The assistance may consist of the individuation of a plan and its presentation to the user. The assistant checks the execution and takes initiative in case of failure, performing a diagnostic analysis and proposing possible alternative strategies. In particular, assistance planning is a distinctive type of interactive planning, in which the planner supports a human being while trying to achieve some complex goal. An intelligent assistant capable of assistance planning decreases the work overload, which is characteristic of several activities. Furthermore, it is of great help to humans working towards the solution of some complex problems [6]. An interactive assistance planner generally has a representation of the task as a partial ordered plan. This feature allows the user to choose among different paths of execution. In addition, it is a hierarchical planner: in this way, action decomposition enables the system to change the repertory of actions used in the plan in order to consider only actions that are executable by the user. The central hypothesis here is that primitive action set can change according on user’s affective state.

We can conceive a prototypical tool in which the assistance to the task-oriented activity is characterized by adaptability to the user’s affective state, through the correlation with the primitive action set. The system would consist of an Interactive Learning Environment (ILE), in order to provide a task domain, and of an assistant to the user’s decision-making [1]. If the system is able to recognize the affective state of the user (in particular, the rate of frustration), it can correlate it with the current set of primitive actions. When a communicated action is not executable, the system decomposes it considering the corresponding node in the action hierarchy and extracting the actions of sub-nodes. The action decomposition can be repeated until it produces a set of actions that the user can execute. The advantage of affective recognition is that, when a change of affective state occurs, the set of primitive actions is automatically updated. This model may evolve differently for each affective state: if a change of his/her mental state is recognized, it may have a corresponding change in primitive action set and consequently in the way in which the task is communicated by the assistant (through the currently executable actions).

If in particular we consider a specific emotion such as frustration, more directly connected to the executive difficulty, the system has to be designed in order to regulate the emotional intensity through action decomposition. When task difficulty is too high for the user, frustration is generated or increased. In turn, frustration affects the cognitive state then reducing the number of executable actions, further increasing stress until impasse occurs. Even if a possible solution could consist of a direct action on the emotional state (e.g. employing humorous or emphatic communication), the planning adaptation itself may perform a regulation of frustration. In fact, after the action decomposition, plan
execution becomes easier and then frustration decreases. In an opposite case (i.e. actions communicated to the user are perceived as too easy), the execution might be boring or annoying, and there is the risk to make the user less motivated to continue the interaction. In this case, some actions can be composed moving up in the hierarchy.

3. Frustration Regulation and Affective Loop

If we suppose to employ assisted task execution as a particular functionality of an ILE, frustration regulation through action decomposition can be considered to be a specific type of affective loop.

Figure 1 shows the case of frustration reduction, the interacting key elements (i.e., difficulty, frustration, and task representation) and the steps of the adaptive planning corresponding to three of the affective loop phases. The steps are described below:

1. **Frustration recognition.** This stage corresponds to the emotional recognition phase of the affective loop, performed with the state-of-the-art sensors.
2. **Action decomposition.** In this phase (corresponding to the response phase of the affective loop) two possible operations are performed. In the first interaction with the user, response consists of the modeling of the primitive actions and the association to the affective state. Action modeling can be performed through the analysis of impasses in the execution. In the following sessions, response consists of the previously modeled set of actions.
3. **Difficulty reduction.** This step corresponds to the emotional elicitation phase of the affective loop. In fact, the presentation of an easier plan to the user reduces the sense of difficulty and frustration decreases.
4. Conclusions

In this paper the potential advantage of affective adaptivity in assisted task execution is proposed. In particular, the focus is on a very simple but general structure (action hierarchy) and mechanism (action decomposition) and the correlation between the set of the immediately executable actions (primitive action set) and the affective state (frustration). The definition of difficulty in terms of depth of the action hierarchy for a given task allows us to distinguish and correlate the objective aspect of task difficulty and the subjective and affective counterpart. Finally, this framework allows us to analyze the mutual interaction between task representation and frustration, and to consider it as specific type of affective loop.

These ideas are not constrained in the domain of task execution, and might be applied to text understanding. In this context, we can consider a concept hierarchy, a set of primitive concepts and a mechanism of concept decomposition that can be adapted to the emotional state of the student. A possible application can be the generation of text in which the terminology is adapted not only to the user knowledge (in particular, the set of primitive concepts) but also to the affective state. If the user is frustrated because the text is difficult to understand, the presentation of content can be modified in order to facilitate reading and understanding.

Finally, assistance to task execution and to text understanding can be integrated in a new generation of ILE in which the frustration recognition and regulation contribute to improve learning.

References


