Analytical Models of Emotions, Learning and Relationships: Towards an Affect-sensitive Cognitive Machine

Barry Kort, Rob Reilly
M.I.T. Media Laboratory
{bkort, reilly}@media.mit.edu

Abstract

Numerous research studies support the claim that affect plays a critical role in decision-making and performance as it influences cognitive processes [see e.g., Damasio, 1994; Goleman, 1995; Picard, 1997]. Despite this body of research the role and function of affect is not generally recognized by the disciplines that address the broad issues of understanding complex systems and complex behavior, especially in the presence of learning. The innovative models and theories that have been proposed to facilitate advancement in the field of human-computer interaction (HCI) tend to focus exclusively on cognitive factors. Consequently, the resulting systems are often unable to adapt to real-world situations in which affective factors play a significant role. We propose several new models for framing a dialogue leading to new insights and innovations that incorporate theories of affect into the design of (affect-sensitive) cognitive machines.

1. Introduction

> Do emotions contribute to intelligence, and if so, what are the implications for the development of a technology of affective computing?
>
> - Robert Provine, *What Questions Are On Psychologist's Minds Today*?

The emerging discipline of Affective Computing has begun to address a variety of research, methodological, and technical issues pertaining to the integration of affect into HCI (e.g., machine recognition of affective states of the user, synthesis of affective states of cartoon avatars or embodied agents, applications incorporating social-emotional intelligence). In order for Affective Computing to become a discipline it should be supported by:

- a novel model that supports model-based reasoning, and,
- an innovative learning cycle model that integrates/accounts for affect.

2. Background

When first introduced, Intelligent Tutoring Systems (ITSs), which were the next generation of Computer Assisted Instruction systems (CAI), “were avowed as the future of education and training” [Jerinic and Devedzic, 2000]. Despite some initial successes (Anderson, 1990; Bonar, 1988; Russell et. al., 1988; Sleeman, 1987; Woolf, 1987), “ITSs have not yet seen general acceptance” [Jerinic and Devedzic, 2000]. And today the ITS community “is still talking about the promise of this technology while searching for the leverage that will encourage its widespread adoption and classroom use” [Jerinic and Devedzic, 2000].

Woolf [1992] observes that typical ITSs consists of:

- a **Domain Knowledge Module**, which contains the information that the tutor is teaching,
- an Expert Module, is more than a mere representation of the data, it is a model of how an expert human teacher would present the Domain Knowledge,
- a **Student Module**, which maintains information that is specific to each user and how far they have progressed,
- a **Tutor/Pedagogical Module**, which is responsible for deciding how and when the domain knowledge is presented; this module emulates the pedagogical approach of an expert teacher (e.g., when to present a new topic, which topic to present, when a review is needed),
- a **Diagnostic/Misconception Module**, which contains the rules used to identify misperceptions, gaps and misunderstandings on the part of the user,
- a **Communication Module**, which is the user interface (e.g., keyboard, mouse, sentic device, screen display/layout). This module answers the question: “How best to present the material to the learner?” and,
- Jerinic and Devedzic [1997] have added an **Explanation Module**, which is “define[d as] the contents of explanations and justifications of the ITS’s learning process, as well as the way they are generated” (e.g., canned text, templates more fully explaining concept *x*, presentation models to explain concept *x* in concert with other modules).

"There is an interplay between emotions and learning, but this interaction is far more complex than previous theories have articulated" [Kort, et al 2001]. It's important to note that the typical ITS largely ignores a user's emotional (affective) state. There is no module to provide for emotion (affective) scaffolding; in a preliminary study Aist, et al. [2002], found support for the contention that emotional (affective) scaffolding can improve the state-of-the-art, at least when provided by a human. It seems apparent that ITSs must learn how to recognize, interpret and react appropriately to a person's affective state.

3. Science and Storymaking

To understand the need for a novel model, let us first examine the current educational model. The current model, as shown in Figure 2, begins with 'data,' which is a collection of answers to questions that the learner has not yet seen fit to ask or needed to ask. Such data becomes 'information' when it answers a question that the learner cares to ask. For the most part, a teacher, who must somehow motivate the student to care enough to seek the answers found in the data, supplies these questions. Studying is like 'panning for gold' where the answers are the 'nuggets' buried in a ton of otherwise uninteresting gravel. Once we have our 'nuggets of information' how do we organize them into a 'body of knowledge'? We may think of 'information' as the pieces of an unassembled jigsaw puzzle, whereas 'knowledge' is the assembled jigsaw puzzle. That is, the question-answer pairs are organized into a coherent structure, in the logical and natural order in which new questions arise as soon as old ones are answered.

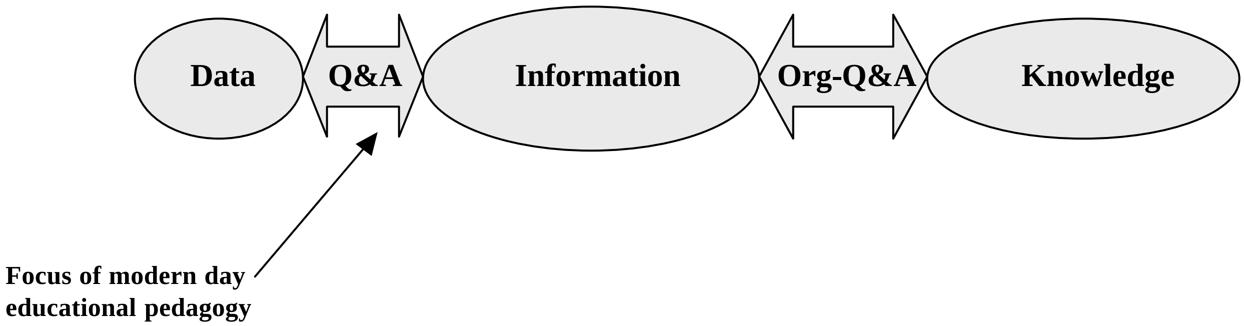

Figure 2 – Old Model: Supports Rule-based Learning

The assembled ‘jigsaw puzzle of knowledge’ reveals a previously hidden picture—a ‘big picture,’ if you will. Or to put it another way, the assembled ‘jigsaw puzzle of knowledge’ is a tapestry into which is woven many otherwise hidden and previously unrevealed stories.

The novel model shown below in Figure 3 goes beyond the current model shown in Figure 2. The focii of attention shifts to the construction of ‘knowledge’ and to the extraction of meaningful ‘insights’ from the ‘big picture.’ When ‘knowledge’ is coupled with a personal or cultural value system, ‘wisdom’ emerges. In other words, wisdom allows us to harness the power of knowledge for beneficial purposes.

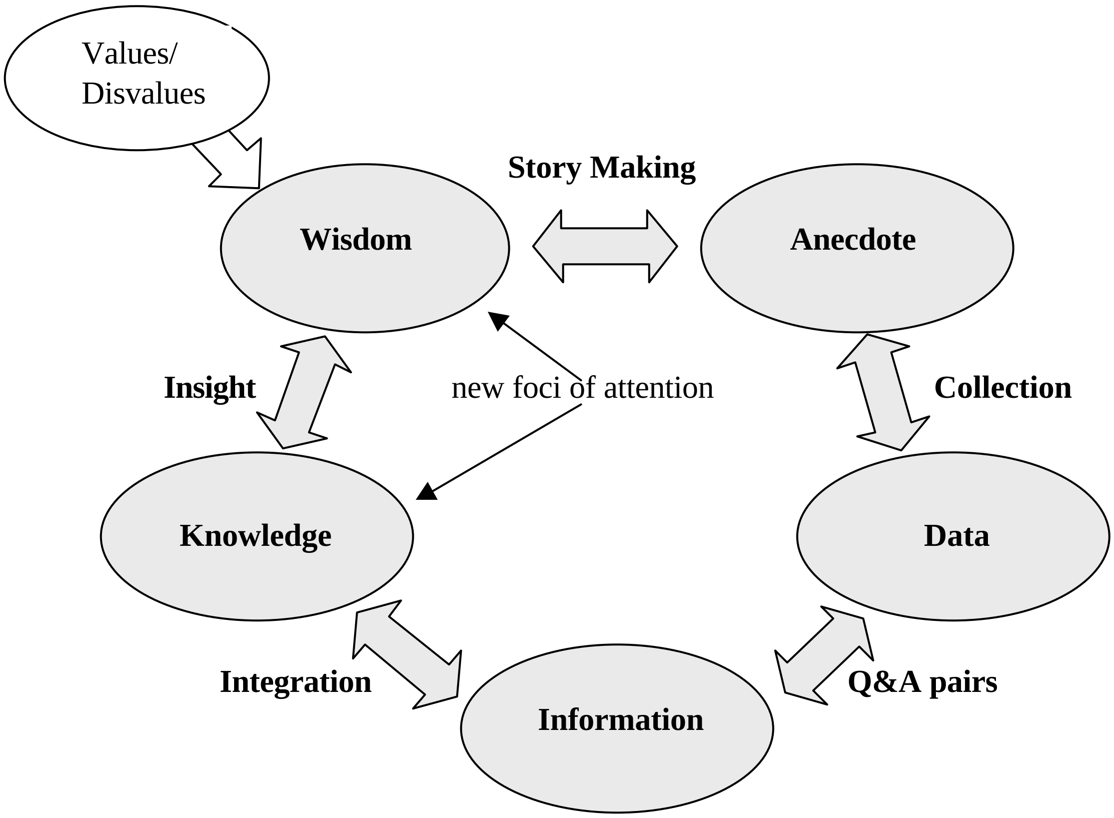

Figure 3 -- New Model: Supports Model-based Reasoning

'Wisdom' affords us the possibility of extracting the stories woven into the tapestry of knowledge. So from 'wisdom' we craft the bardic arts of story making and story telling. The ancients crafted myths and legends. These were the prototypical stories of their cultures, which were intended to impart 'wisdom.' A story is thus an anecdote drawn from the culture. A well-crafted anecdote or story has value both as an amusement and as a source of insight into the world from which it is drawn. And the plural of 'anecdote' is data—a collection of anecdotal stories or evidence. This observation closes the loop in Figure 3.

Figure 3 suggests a novel model that, on a fundamental level, supports an improved educational pedagogy. This will serve as a foundation for the next part of our model—how a learner's affective state should be incorporated into the overall model.

4. Models of Emotions and Learning

> The extent to which emotional upsets can interfere with mental life is no news to teachers. Students who are anxious, angry, or depressed don't learn; people who are caught in these states do not take in information efficiently or deal with it well.
>
> - Daniel Goleman, *Emotional Intelligence*

In an attempt to install/build/re-engineer the current state of educational pedagogy, educators should first look to expert teachers who are adept at recognizing the emotional state of learners, and, based upon their observations, take some action that scaffolds learning in a positive manner. But what do these expert teachers *see* and how do they decide upon a course of action? How do students who have strayed from *learning* return to a productive path, such as the one that Csikszentmihalyi [1990] refers to as the "zone of flow"? This notion that a student's affective (emotional) state impacts learning and that appropriate intervention based upon that affective state would facilitate learning is the concept that we propose to explore in-depth.

To prove our point, note that skilled humans can assess emotional signals with varying degrees of precision. For example, researchers are beginning to make progress giving computers similar abilities to accurately recognize affective expressions [Picard, 2000; Scheirer, et. al., 1999], facial expressions [Bartlett, 1999; Cohn, et al., 1999; Donato, 1999; DeSilva, 1997; Ekman, 1997; Essa, 1995], and gestural expression [Chen, et al., 1998; Huang, 1998]. Although computers only perform as well as people in highly restricted domains, we believe that:

?? accurately identifying a learner's cognitive-emotive state is a critical observation that will enable teachers to provide learners with an efficient and pleasurable learning experience, and,

?? unobtrusive highly accurate technology will be developed to accurately assess actions in less restricted domains (see e.g., Kapoor, et al., 2001).

Our own preliminary pilot studies with elementary school children suggest that a human observer can assess the affective emotional state of a student with reasonable reliability based on observation of facial expressions, gross body language, and the content and tone of speech. If the human observer is also acting in the role of coach or mentor, these assessments can be confirmed or refined by direct conversation (e.g. simply asking the student if she is confused or frustrated before offering to provide coaching or hints). Moreover, successful learning is frequently marked by an unmistakable elation, often jointly celebrated with "high fives." In some cases, the "Aha!" moment is so dramatic, it verges on the epiphanetic. One of the great joys for an educator is to bring a student to such a moment of triumph. But how can computers acquire this same level of proficiency as that of gifted coaches, mentors, and teachers?

Our first step is to offer a model of a learning cycle, which integrates affect. Figure 4 suggests six possible emotion axes that may arise in the course of learning. Figures 5a and 5b interweave the emotion axes shown in Figure 4 with the cognitive dynamics of the learning process. In Figure 5, the positive valence (more pleasurable) emotions are on the right; the negative valence (more unpleasant) emotions are on the left. The vertical axis is what we call the Learning Axis, and symbolizes the construction of knowledge upward, and the discarding of misconceptions downward.

Axis	-1. 0	-0. 5		0	+0. 5	+1. 0
Anxiety-Confidence	Anxiety	Worry	Discomfort	Comfort	Hopefulness	Confidence
Ennui-Fascination	Ennui	Boredom	Indifference	Interest	Curiosity	Fascination
Frustration-Euphoria	Frustration	Puzzlement	Confusion	Insight	Enlightenment	Euphoria
Dispirited-Enthusiasm	Dispirited	Disappointed	Dissatisfied	Satisfied	Thrilled	Enthusiasm
Terror -Excitement	Terror	Dread	Apprehension	Calm	Anticipatory	Excitement
Humiliated-Proud	Humiliated	Embarrassed	Self-conscious	Pleased	Satisfied	Proud

Figure 4 – Emotion sets possibly relevant to learning

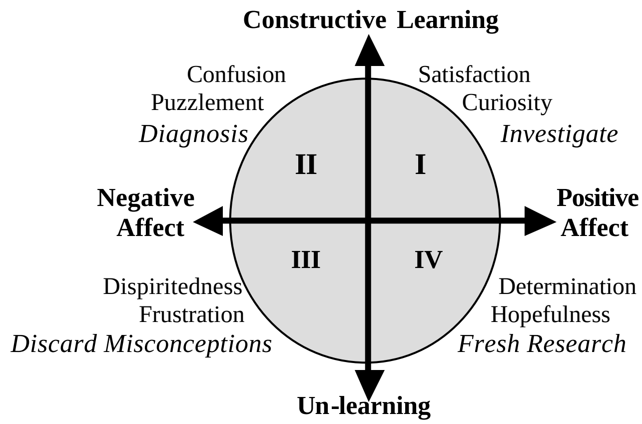

Figure 5a – Four Quadrant model relating phases of learning to emotions in Figure 4

Students ideally begin in Quadrant I or II: they might be curious or fascinated about a new topic of interest (Quadrant I) or they might be puzzled and motivated to reduce confusion (Quadrant II). In either case, they are in the top half of the space if their focus is on constructing or testing knowledge. Movement happens in this space as learning proceeds. For example, when solving a puzzle in *The Incredible Machine*, a student gets a bright idea how to implement a solution and then builds its simulation. If she runs the simulation and it fails, she sees that her idea has some part that doesn't work—that needs to be diagnosed and reconstructed. At this point the student may move down into the lower half of the diagram (Quadrant III) into the 'dark teatime of the soul' while discarding misconceptions and unproductive ideas. As she consolidates her knowledge—what works and what does not—with awareness of a sense of making progress, she advances to Quadrant IV. Getting another fresh idea propels the student back into the upper half of the space (Quadrant I). Thus, a typical learning experience involves a range of emotions, cycling the student around the four quadrant cognitive-emotive space as they learn.

If one visualizes a version of Figure 5a (and Figure 5b) for each axis in Figure 4, then at any given instant, the student might be in multiple Quadrants with respect to different axes. They might be in Quadrant II with respect to feeling frustrated and simultaneously in Quadrant I with respect to interest level. It is important to recognize that a range of emotions occurs naturally in a real learning process, and it is not simply the case that the positive emotions are the good ones.

We do not foresee trying to keep the student in Quadrant I, but rather to help him see that the cyclic nature is natural in learning science, mathematics, engineering or technology (SMET), and that when he lands in the negative half, it is an inevitable part of the cycle. Our aim is to help students to keep orbiting the loop, teaching them to propel themselves, especially after a setback.

A third axis (not shown) can be envisioned as extending out of the plane of the page—the cumulative knowledge axis. If one visualizes the above dynamics of moving from Quadrant I to II to III to IV as an orbit, then, when this third dimension is added, one obtains an excelsior spiral. In Quadrant I, anticipation and expectation are high, as the learner builds ideas and concepts and tries them out. Emotional mood decays over time either from boredom or from disappointment. In Quadrant II, the rate of construction of working knowledge diminishes, and negative emotions emerge as progress wanes. In Quadrant III, as the negative affect runs its course, the learner discards misconceptions and ideas that didn't pan out. In Quadrant IV, the

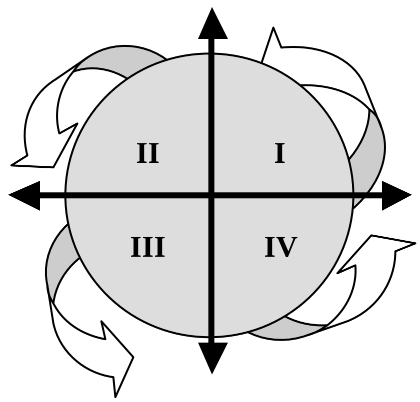

Figure 5b – Circular and helical flow of emotion in Four Quadrant model

learner recovers hopefulness and positive attitude as the knowledge set is now cleared of unworkable and unproductive concepts, and the cycle begins anew. In building a complete and correct mental model associated with a learning opportunity, the learner may experience multiple cycles until completion of the learning exercise. Note that the orbit doesn't close on itself, but gradually spirals around the cumulative knowledge axis.

We are in the process of performing empirical research on this model. We have conducted several pilot research projects, which appear to confirm the model. (Note: Interested readers can find more about this work in our reference list.)

5. Conclusion

Our models are inspired by theory often used to describe complex dynamic interactions in engineering systems. As such, they are not intended to explain how learning works, but rather to provide a framework for thinking and posing questions about the role of emotions in learning. As with any metaphor, the model has its limits. The model does not encompass all aspects of the complex interaction between emotions and learning, but begins to describe some of the key phenomena that needs to be considered in metacognition.

These models go beyond previous research studies not just in the range of emotions addressed, but also in an attempt to formalize an analytical model that describes the dynamics of a learner's emotional states, and does so in a language that supports metacognitive analysis.

6. Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 0087768. Any opinions, findings, or conclusions or recommendations expressed in this material are those of the author(s) and does not necessarily reflect the views of the National Science Foundation.

7. References

BARTLETT, M., J.C. HAGER, P. EKMAN, and T. SEJNOWSKI. (1999). Measuring Facial Expression by Computer Image Analysis. *Psychophysiology*, vol. 36, pp. 253-263.

BRANSFORD, JOHN, ANN L. BROWN, and RODNEY COCKING (Eds.) (1999). *How People Learn: Brain, Mind, Experience, and School*. Washington DC: National Academy Press.

CHEN, L.S., T.S. HUANG, T. MIYASATO, and R. NAKATSU. (1998). Multimodal Human Emotion/Expression Recognition. *Proceedings of 3rd International Conference on Automated Face and Gesture Recognition*, pp366-371.

COHN, J.F., A.J. ZLOCHOWER, J. LIEN, and T. KANADE. (1999). Automated Face Analysis by Feature Point Tracking has High Concurrent Validity with Manual FACS Coding, *Psychophysiology*, vol. 36, pp35-43.

CSIKSZENTMIHALYI, M. (1990). *Flow: The Psychology of Optimal Experience*, Harper-Row: NY.

DAMASIO, A.R., (1994). *Descartes Error: Emotion, Reason and the Human Brain*, G.P. Putnam Sons: NY.

DEL SOLDATO, T. (1994). *Motivation in Tutoring Systems*. Tech. Rep. CSRP 303, School of Cognitive and Computing Science, The University of Sussex, UK.

DESILVA, L.C., T. MIYASATO, and R. NAKATSU, Facial emotion recognition using multi-modal information, in *Proc. IEEE Int. Conf. on Info., Comm. and Sig. Proc.*, (Singapore), pp. 397-401, Sept 1997.

DE VINCENTE, A. and PAIN, H. Motivation Self-Report in ITS. In Lajoie, S. P. and Vivet, M. eds. *Proc. of the Ninth World Conference on Artificial Intelligence in Education*, pp. 651-653, Amsterdam, IOS Press

DONATO, G., M.S. BARTLETT, J.C. HAGER, P. EKMAN, and T.J. SEJNOWSKI, Classifying facial actions, *IEEE Pattern Analy. and Mach. Intell.*, vol. 21, pp. 974--989, October 1999.

EKMAN, PAUL., (1997). *Facial Action Coding System*, Consulting Psychologists Press.

ESSA, I. and PENTLAND, A., Coding, analysis, interpretation and recognition of facial expressions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 757--763, July 1997.

GOLDMAN, DANIEL., (1995). *Emotional Intelligence*. Bantam Books: New York.

HARO, A., ESSA, I., and FLICKNER, M., Detecting and Tracking Eyes by Using their Physiological Properties, Dynamics and Appearance, In *Proceedings of IEEE Computer Vision and Pattern Recognition*, Hilton Head, SC, June 2000.

HUANG, T.S., L.S. CHEN, and H. TAO. (1998). Bimodal Emotion Recognition by Man and Machine. *ATR Workshop on Virtual Communication Environments*.

KAPOOR, ASHISH, SELENE MOTA and ROSALIND PICARD. (2001). Towards a Learning Companion that Recognizes Affect, *Proceedings of AAAI 2001*.

KLEIN, J. (1999), *Computer Response to User Frustration*, Master's thesis, MIT Media Lab.

MATSUBARA, Y., NAGAMACHI, M., (1996). Motivation systems and motivation models for intelligent tutoring, in Claude Frasson, et. al., (Eds.), *Proceedings of the Third International Conference in Intelligent Tutoring Systems*, 1996, pp. 139-147.

OLSON, R.K. and WISE, B. (1987). Computer Speech in Reading Instruction, In D Reinking (Ed.). *Computers and Reading: Issues in Theory and Practice*. New York: Teachers College Press.

PAPERT, SEYMOUR (1993). *The Children's Machine: Rethinking School in the Age of the Computer*, Basic Books: New York.

PIAGET, JEAN. (1952). *The Origins of Intelligence in Children*. M. Cook translator. New York: International Universities Press.

PICARD, ROSALIND W. (2000). *Toward Computers that Recognize and Respond to User Emotions*, IBM Systems Journal, Vol. 39 (3 and 4), p. 705.

PICARD, ROSALIND W. (1997). *Affective Computing*. MIT Press: Cambridge, MA.

PROVINE, ROBERT, (1998). *What Questions Are On Psychologist's Minds Today*? Available on-line as of August 1, 2001 at: http://www.edge.org/3rd_culture/myers/index.html

RICH, C., WATERS, R. C., STROHECKER, C., SCHABES, Y., FREMEN, W. T., TORANCE, M. C., GOLDING, A., ROTH, M. (1994), *A Prototype Interactive Environment for Collaboration and Learning*, Technical Report TR-94-06, http://www.merl.com/projects/emp/index.html

SCHEIRER, J., R. FERNANDEZ and ROSALIND. W. PICARD (1999), Expression Glasses: A Wearable Device for Facial Expression Recognition, *Proceedings of CHI*, February 1999.

YACOOB, Y. and DAVIS, L., Recognizing human facial expressions from log image sequences using optical flow, *IEEE Transaction on Pattern Analysis and Machine Intell*igence, vol. 18, pp.636-642, June 1996.

YINGLI TIAN, KANADE, T. and COHN, J. F., Recognizing Action Units for Facial Expression Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, February, 2001.