

Personality, Affect and Emotion Taxonomy for Socially Intelligent Agents

Christine L. Lisetti
 Department of Computer Science
 University of Central Florida
 Orlando, FL 32816
 lisetti@cs.ucf.edu

Abstract

In this article, we describe an Affective Knowledge Representation (AKR) scheme to represent emotion schemata to be used in the design a variety of socially intelligent artificial agents. Our approach in this article and in the applications of our AKR scheme, focusses on the notion of “social expertise” of socially intelligent agents in terms of their 1) external behavior and 2) internal motivational goal-based abilities. AKR is derived from combining multiple emotion theories in a useful hierarchical model of affective phenomena and includes a taxonomy of affect, mood, emotion, and personality, as well as a framework for emotional state dynamics. Our model is being applied to design and implement two systems: 1) EBA, an Emotion-Based Architecture for two autonomous robots, 2) MAUI, a Multimodal Affective User Interface agent.

Affective Knowledge Representation (AKR)

The functional role of emotions has been recently fully recognized as essential for intelligent agents with limited resources operating in an complex and unpredictable environment (Murphy, Lisetti et al. 2002). In order to contribute to rendering artificial intelligent agents socially more competent, we combined and reconciled aspects of the main current theories of affect, mood and emotion (Ortony et al. 1988), (Frijda, 1986), (Wierzbicka, 1992), into a simplified comprehensive, (but not complete) taxonomy of affect, mood and emotion for computational Affective Knowledge Representation (AKR).

We created AKR to enable the design of a variety of artificial autonomous (i.e. self-motivated) socially competent agents, from robotics (Murphy et al. 2002), (Lisetti, 1997), to user-modeling (Lisetti and Bianchi 2002), to human-computer interaction (Hayes-Roth et al. 1998), to multi-agent systems and distributed AI.¹ Our taxonomy of affective states in Figure 1 is aimed at differentiating among the variety of affective states by using values of well-defined componential attributes. First we define our use of terms throughout our work.

¹Copyright 2002, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

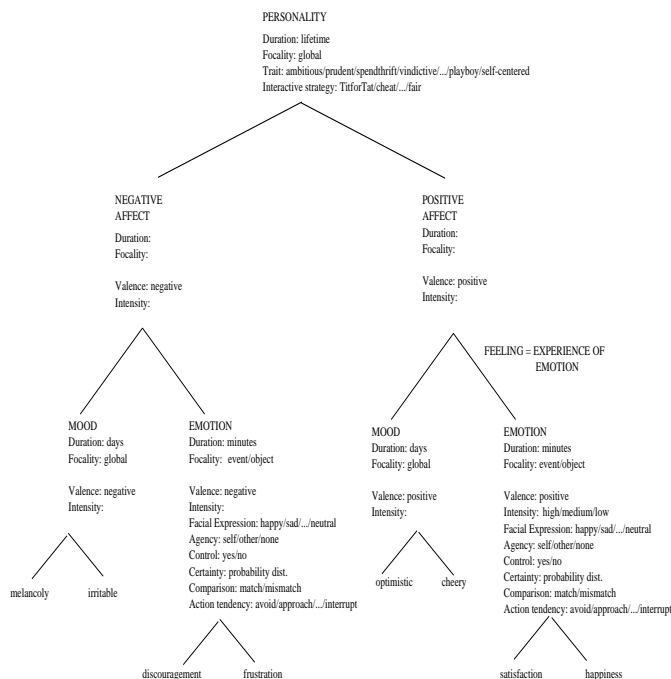


Figure 1: Hierarchical Model of Personality, Affect, Mood and Emotion

Personality: We identify *personality* as representing characteristics of an autonomous organism that account for consistently chosen patterns of mental reaction including behavior, emotions and thoughts over situations and time (Moffat, 1997).

Affect: Affect varies along two dimensions: (i) *valence* which can be positive or negative (the pleasant and unpleasant dimension) and (ii) *intensity* which varies in terms of degree. Affect considered as a state is therefore a coarse-grained phenomena (it can also be considered as a phenomena including finer-grained states as shown in Figure 1).

Mood: Moods are affective phenomena encoding coarser-grained information and of shorter duration

than emotions.

Emotion: We identify *emotions as changes in activation of behavioral dispositions or transformation of dispositions to act, caused by relevant events or circumstances.*

Because emotions are at the bottom of the hierarchical model, emotions do not necessarily imply personalities, since some emotions might be experienced by different types of agents (artificial or natural). The type of personality, is inherited from the higher node in the tree, allowing agents of different personality type to still experience the full range of possible emotions as advocated by other computational approaches (Castelfranchi, 1997). Because we adopt the functional view of emotions, which identifies emotions as related to goals and action tendencies (such as self-preservation, avoid negative experiences, approach enjoyable things, etc.), our model is compatible with goal-oriented theory of personality (Carbognell 1980).

In addition, because the interactive strategies (Tit-for-Tat, cheat, etc.) and preferences (avoid/dislike negative and seek/like positive affect) are specified in the model at a higher level than at the emotion level, and because personality explicitly is represented as lasting a lifetime and not related with any specific event, this approach is in agreement with other views which emphasize the main distinction between personality (stable) and emotion (changeable). Positive and negative affect can be equivalent to positive and negative attitude, while personality traits are one component of the entire personality description). This approach is also consistent with the idea that different personalities can influence an agent's propensity toward a particular set of emotions and moods.

Emotion Components

In order to address some of the difficulties of the previous computational approaches pointed out by Pfeifer (Pfeifer 1988) – namely the lack of representation of physiological and subjective parameters – we do not split 'emotion' and 'cognition', but rather merge them into a structure that encapsulates simultaneously each of the three phenomena accompanying emotions: (1) ANS arousal, signaling affect intensity and valence; (2) expression, for now only facial expression (but could also include vocal, and body posture); and (3) subjective experience, including cognitive appraisals (such as modal beliefs, criteria, standards, etc.).

In an effort to identify what makes one emotion different from another, we include elements from the "cognitive view" of emotions, which advocates a componential approach (Leventhal, 1987), (Frijda, 1986), (Ortony, 1988) (Roseman, 1996). From this approach, cognitive structures associated with emotions are considered to represent the subject's checks (appraisal or evaluation) of the events confronting them.

These checks are part of the subjective experience of emotion and can be represented with a limited number of components. Each type of checks is described

as a unique pattern of such components, or dimension values. As with the set of basic emotions which varies among theories, several dimensions are often considered, but the following are found in most analyses: valence (pleasantness or unpleasantness), intensity/urgency, agency/responsibility, novelty, controllability, modifiability, certainty, external social norms compatibility, and internal standards compatibility. We also included duration and focality which differentiate emotions from moods, for future potential expansion of the model. These components are indicated shown below:

- **facial expression:** *happy/sad/surprised/disgusted/angry/fearful/neutral* is used to store the facial expression associated with the emotion. Some emotions are not associated with any specific facial expression (neutral), or can vary among cultures and individuals.
- **valence:** *positive/negative* is used to describe the pleasant or unpleasant dimension of an affective state. Each affective phenomena is associated with a valence, except for the emotion of surprise which can be either positive or negative depending upon the nature of the stimulus.
- **intensity:** *very high/high/medium/low/very low*. varies in terms of degree. The intensity of an affective state is relevant to the importance, relevance and urgency of the message that the state carries²
- **duration:** *lifetime/days/minutes* is used to indicate that moods are more muted and last longer than emotions, which is indicated by the *duration* attribute measured in terms of days, as opposed to minutes in the case of emotions; it can also be used to resolve the conflict between personality and emotion by assuming that the underlying mechanisms are essentially the same, and that only the time-course and temporal consistency of their influence varies: personalities can be permanent and last a lifetime.
- **focality:** *global/event/object* is used to indicate whether the affective phenomena is global (the cause may not be a meaningful event but rather a biochemical change) as in moods in which the cause has become detached from the felt action readiness (the cause may not be an experienced meaningful event, but it may be biochemical), or on the other hand, as in emotions which are mostly about something: an event (the trigger to surprise) or an object (the object of jealousy). Globality can also differentiate emotions: depression from sadness, bliss from joy, anxiety from fear. In depression the world as a whole appears devoid of intentional objects; similarly in happiness, the environment as a whole appears tinted with positive valence.

²In natural organisms, valence and intensity is signaled by the activity of the autonomic nervous system, along the physiological dimension generated by the body proper, and do not necessarily involve the cognitive apparatus.

- **agency:** *self/other* is used to indicate who was responsible for the emotion, the agent itself *self*, or someone else *other*. For example, if the agent is angry at someone, the agent parameter will point to that person; but if the agent is depressed, agency is most likely to point to self.
- **novelty:** *match/mismatch* is used to refer to whether a novel and unexpected stimulus occurred causing mismatch with the subject’s expectations regarding the stimulus triggered.
- **intentionality:** *other/self* is used to refer to whether the triggering event is perceived as caused by some live intending agent. In anger, it is other, whereas in self-hatred and guilt it is self.
- **controllability:** *high, medium, low, none* is used to refer to how much the subject believes s/he/it can control the current situation. Controllability is the component that turns danger from threat into challenges, and therefore negative into positive emotion. Change from angry protest to despair and resignation can be interpreted as a consequence of the fact that uncontrollability gradually dawns.
- **modifiability:** *high, medium, low, none* is used to refer to duration and time perspective, or to the judgment that a course of events is capable of changing. Modifiability carries with it the past, in the sense that what has been for a long time may well be forever. It can also apply to current events, e.g. suffering a situation as if it will never end, feeling of self-confidence.
- **certainty:** *certain, noncertain, non-uncertain* is used to refer to anticipation of effects to come, and how (subjectively) certain the subject is about the consequences of the situation. For example, joy implies absence of uncertainty (uncertainty about how a friend will respond takes away the joy of going to meet them), yet the aspect of certainty is implicit (hence our three values).
- **legitimacy:** *yes,no* is used to indicate whether the emotion is experienced as a legitimate states.
- **external (social) norm:** *compatible/uncompatible* is used to refer to whether the event (usually an action) conforms to social norms, cultural conventions, or expectations of significant other.
- **internal (self) standard:** *compatible/uncompatible* is used to refer to whether the event (usually an action) is consistent with internalized personal standards as part of the self concept or ideal self.
- **action tendency:** identifies the most appropriate (suite of) actions to be taken from that emotional state.
- **causal chain:** identifies the causation of a stimulus event (described next).

Functional Attributes and Action Tendencies

From the Darwinian categorical theory of emotions, emotions can be discretely categorized. Emotions are considered as mental and physiological processes, caused by the perception of general categories of event, that elicits internal and external signals and matching suite of action plans. This Darwinian perspective proposes that bridging the gaps of rationality becomes possible if many specific emotional states are mapped into a few broad classes of reaction, or *action tendencies*.

Action tendency: Emotions which are called “primary” or “basic” are such in the sense that they are considered to correspond to distinct and elementary forms of action tendency. Each “discrete emotion” calls into readiness a small and distinctive suite of action plans that has been selected as appropriate when in the current emotional state. Thus in broadly defined recurring circumstances that are relevant to goals, each emotion prompts both the individual and the group in a way that has been evolutionarily more successful than alternative kinds of prompting.

The number and choice of what is called basic or primary emotions vary among various emotion theories, and we have selected the ones that seem to reoccur consistently across emotion theories. Their associated action tendency are listed in the Table 1.

ACTION TENDENCY	END STATE	FUNCTION	EMOTION
Approach	Access	Permit consumption	Desire
Avoid	Own inaccessibility	Protection	Fear
Attend	Identification	Orientation	Interest
Reject	Removal of object	Protection	Disgust
Agnostic	Removal of obstruction	Regain of Control	Anger
Interrupt	Reorientation	Reorientation	Surprise
Free activate	action tendency’s end state	Generalized Readiness	Joy
Inactivity	action tendency’s end state	Recuperation	Content
Inhibit	absence of response	Caution	Anxiety
Prepare			

Table 1: Action Tendency Table

Emotion Beliefs and Causal Chains

We adapted the semantic meta-definitions of emotion concepts developed by Wierzbicka using language independent primitives (Wierzbicka, 1992) to create the *causal chain*.

Causal chain: A causal chain of events describes the *subjective cognitive experience* components which are associated with the emotion, the beliefs, the goals, and their achievement or lack of. Illustrations of these are found at the bottom of each example in Tables 2 and 3. These can be spoken via speech synthesis for the agent to express and describe the cognitive interpretation of its state.

HAPPINESS
Emotion Components:
Agent = <i>RobotWaiter</i>
Facial Expression = <i>happy</i>
Valence = <i>positive</i>
Intensity = <i>high</i>
Duration = <i>minutes</i>
Focality = <i>event</i>
Agency = <i>unspec.</i>
Novelty = <i>match</i>
Controllability = <i>unspec.</i>
Modifiability = <i>high</i>
Certainty = <i>non-uncertain</i>
External Norm = <i>unspec.</i>
Internal Standard = <i>unpec.</i>
Action Tendency = <i>FreeActivate</i>
Body Posture = <i>relaxed</i>
Vocal Tone = <i>neutral</i>
Context = <i>Serving</i>
Causal Chain:
something very good is happening
I want this
because of this I feel very good

Table 2: Emotion schema for *HAPPY*

FRUSTRATION
Emotion Components:
Agent = <i>RobotAssistant</i>
Facial Expression = <i>angry</i>
Valence = <i>negative</i>
Intensity = <i>medium</i>
Duration = <i>minutes</i>
Focality = <i>object</i>
Agency = <i>other</i>
Novelty = <i>mismatch</i>
Controllability = <i>none</i>
Modifiability = <i>medium</i>
Certainty = <i>non certain</i>
External Norm = <i>unspec.</i>
Internal Standard = <i>unpec.</i>
Action Tendency = <i>ChangeStrategy</i>
Body Posture = <i>agitated</i>
Vocal Tone = <i>high</i>
Context = <i>Serving</i>
Causal Chain:
I want to do something
I can't do this
because of this I feel bad

Table 3: Emotion schema for *FRUSTRATION*

Emotional States Dynamics

Our framework also establishes a dynamic model of emotional states to: 1) generate emotional states for *autonomous agents* given the agent’s current emotional state and the current set of contextual inputs it receives (which can either be external or internal as explained below); 2) help infer and/or predict the user’s next emotional state for *multimodal affective interface agents* given the user’s current state and the contextual situation.

External Event as Inputs

For example, in Figure 2 we present an example model of emotions’ dynamics. It contains four emotional states: *HAPPY*, *CONCERNED*, *FRUSTRATED* and *ANGRY* (with *NEUTRAL* as the starting point). The transitions among the states are caused by environmental inputs or responses of the system, and they are divided into categories of positive progress toward goal and negative progress. Using this dynamic model we can predict that an agent that is in a *HAPPY*

state will remain *HAPPY* given positive inputs, will become *CONCERNED* given negative inputs toward its goal (e.g. obstacles of some sort dependent on the context), increasingly so until it reaches the *FRUSTRATED* state. At which point it can receive positive inputs and move back to a *CONCERNED* or *HAPPY* state (depending on the importance of the positive input).

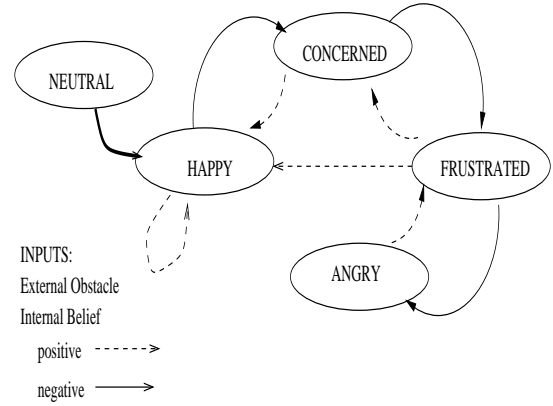


Figure 2: Emotional State Dynamics: Determined Personality

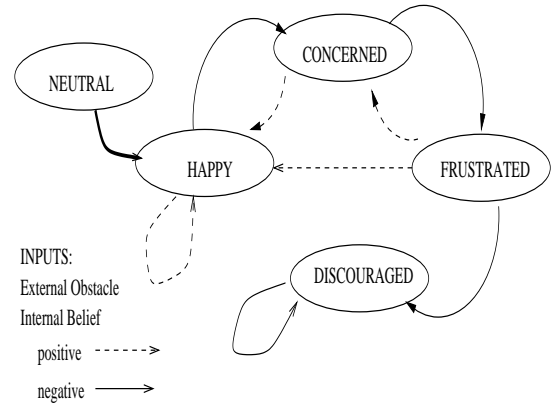


Figure 3: Emotional State Dynamics: Meek Personality

Internal Beliefs as Inputs

An individual’s emotions can change in regard to an event, and these changes may be the result of their own efforts, not simply the result of an independent process directed by external events or social rules. Emotional changes indeed occur as a result of a number of processes, involving emotion dynamics rather than simply outside circumstances or the force of culture.

A simple example is shown in Figure 3, where a negative internal belief such as “I can’t do this” would

keep the agent in its current DISCOURAGED state forever. Should the agent manage to change its internal belief for a positive input (e.g. “I can indeed do this”), it would switch to a state of HOPEFULNESS (not shown). Other examples of such internal self-adjustments abound.

For example, one may experience being guilty about being angry, or being depressed about feeling responsible. Emotional patterns can be transformed and change as a result of circumstances, which in turn provoke further emotions. In our example, when and if one realizes that the experience of being angry was indeed justified, the feeling of being guilty about being angry vanishes. Similarly the feeling of being depressed about feeling responsible would vanish with the understanding of one’s true lack of responsibility.

The first change would occur when one specific component of the anger emotion, namely *legitimacy*, is updated from its previous negative value to a positive one. The second change would occur when one updates the *agency* attribute of the responsible feeling from self to other. Another example of such a displacement of current emotions is found when love turns to duty. A MAUI agent aware of the user’s state can assist the user to move from a negative to a positive one by pointing what an adjustment of parameter (i.e. belief) can do.

In Summary: In our taxonomy, each emotion is considered as a collection of emotion components, such as its valence (the pleasant or unpleasant dimension), its intensity (mild, high, extreme), etc. In our representation, we also included the action tendency of each emotion (Frijda, 1986) which corresponds to the signal that the emotional state experienced points to: a small and distinctive suite of action plans that has been (evolutionarily) selected as appropriate, e.g. approach, avoid, reject, continue, change strategy, etc.

Implementations and Applications

(EBA) Emotion-Based Architecture for Autonomous Robots: We have implemented a script version – adumbrated earlier in (Lisetti, 1997) – of one layer of the model on two autonomous robots who competed at that AAI Mobile Robot Competition and won the Nils Nilsson Award for Integrating AI Technologies (Murphy, Lisetti et al., 2002).

(MAUI) MultiModal Affective User Interface Agents: We have used our model to build Models Of the User’s Emotions (our MOUE system) during interaction discussed in (Lisetti and Bianchi, 2002). We describe the MOUE application for telemedicine in (Lisetti et al. 2001), and show the MAUI agent with additional affective modalities in (Lisetti and Nasoz, 2002).

Game Theoretic Agents: The taxonomy is also being used to guide the decision-theoretic formalization of the role of emotions role in a simple decision-making game theoretic framework (Lisetti and Gmytrasiewicz, 2000).

Acknowledgement

The author would like to acknowledge *Intel Corporation*, *Interval Research Corporation* and the *Office of Naval Research* for partial funding for this research.

References

- Carbonell, 1980. Towards a Process Model of Personality Traits. *Artificial Intelligence*, 15:49-74.
- Castelfranchi, C. De Rosis, F. and Falcone, R. Social Attitudes and Personalities in Agents. Menlo Par, CA: AAAI Press, 1997.
- Frijda, N. 1986. *The Emotions*. New-York: Cambridge University Press.
- Hayes-Roth, B., Ball, G. and Lisetti, C. and Picard, R. and Stern, A. Panel of Emotions in the Interface. In *Proc. of IUI’98*.
- Leventhal, H. and Scherer, K. 1987. The Relationship of Emotion to Cognition: A Functional Approach to a Semantic Controversy. *Cognition and Emotion*, 1(1) 3-28.
- Lisetti, C. and Nasoz, F. 2002. Application of the Multilevel Process Theory of Emotion to User-Modeling. In *Proc. of FLAIRS’02*. Menlo Park, CA: AAAI Press.
- Lisetti, C. and Bianchi, N. 2002. Modeling Multimodal Expression of User’s Affective Subjective Experience. *User Modeling and User-Adapted Interaction*, 12(1).
- Lisetti, C. and Douglas, M. and LeRouge, C. 2001. Intelligent Affective Interfaces: A User-Modeling Approach for Telemedicine. In *Proceedings of the First International Conference on Universal Access in Human-Computer Interaction*. Elsevier Science Publisher.
- Lisetti, C. and Gmytrasiewicz, P. 2000. Decisions, Decisions... and the Role of Emotions in the Process. In *Proceedings of the AAAI Fall Symposium Series on Socially Intelligent Agents: The Human in the Loop (also AAAI Technical Report FS-00-04)*. Menlo Park, CA: AAAI Press.
- Lisetti, C. 1997. Motives for Intelligent Agents: Computational Scripts for Emotion Concepts. In *Proceedings of Sixth Scandinavian Conference on Artificial Intelligence*, 59-70. Amsterdam, Netherlands: IOS Press Frontiers in Artificial Intelligence and Applications.
- Ortony, A., Clore, G. and Collins, 1988. *The Cognitive Structure of Emotions*. New-York: Cambridge University Press.
- Moffat, D. 1997. Personality Parameters and Programs. In *Creating Personalities for Synthetic Actors*, Trappl (Ed.), Springer.
- Murphy, R. and Lisetti, C. and Irish, L. and Tardif, R. and Gage, A. (in press). Emotion-Based Control of Cooperating Heterogeneous Mobile Robots. *IEEE Transactions on Robotics and Automation*.
- Wierzbicka, A. 1992. Defining Emotion Concepts. *Cognitive Science* 16: 539-581.