Mixed Emotion Modeling

Valeria Carofiglio, Fiorella de Rosis and Roberto Grassano

Department of Informatics
University of Bari
Via Orabona 4, 70126 Bari, Italy, +39 080 5443284
{derosis, carofiglio, grassano@di.uniba.it}
http://aos2.uniba.it:8080/IntInt.html

In the 'Pure vs Mixed Emotions' Section of her book on 'Affective Computing', Rose Picard introduces the following example: after winning a Marathon, a professional runner described *"feeling tremendously **happy** for winning the race, **surprised** because she believed she would not win, somewhat **sad** that the race was over and a bit **fearful** because during the race she had acute abdominal pain"* (Picard, 2000). A runner's friend, when assisting to that Marathon, was probably *"**happy for her** because she won the race, although a bit **envious** for not being able to participate to it and **sorry for** seing her so tired, at the end"*.

How is it that these two persons referred to feel this different mixture of emotions? Clearly, the main source of difference is due to the different structure of beliefs and goals of their minds. In the runner, the *intensity* of **fear** during the race was probably related, at the same time, to the *importance* she assigned to her goal of winning it and to variations in the *probability* of achieving this goal, that she dynamically revised during the race. The importance of this goal also affected the intensity of **happiness** (or **satisfaction**) of achieving it, while **surprise** was probably due to a difference between the likelihood she attached to achieving the goal at the beginning of the race and the final result. The **sadness** that the event was over might be a mixed emotion in its turn, some combination of **nostalgia** for a pleasant past event and **hope** to be again in a similar situation, in the future. The mixing of emotions in the runner's friend was probably due to a mixing of goals of approximately equal weight: **happy-for** is due her desire of achieving 'the good of her friend' (conditioned to her winning the race) **sorry-for** to her desire of 'preserving her from bad' (illness, in this case) and **envy** to her desire of 'dominating her', in a way.

So: differences between the two persons, in the example, are due to differences in the beliefs, the goals they want to achieve, the weights they assign to achieving them and the structure of links between beliefs and goals; variations of these measures with time seem to govern cognitively-generated emotions. This kind of cognitive emotion modeling is supported by other authors (for instance, Oatley and Johnson Laird, 1987). We claim that *Dynamic Belief Networks* are a mathematical tool that enables representing appropriately the dynamic arousal, evolution and disappearing of emotions, and the way these phenomena are related to personality factors and to the social context. We illustrated in detail the logics behind this modeling method in another paper, in which we discussed the advantages it offers in driving the 'affective' behavior of an Embodied Conversational Agent that we are building in the scope of the European Project 'Magicster'[1] (de Rosis et al, submitted). In this position paper, we summarise its main features, to focus our description on how the two metaphors of mixing emotions proposed by Rose Picard may be modeled with this formalism. We will restrict our analysis to goal-based emotions, in Ortony's classification (Ortony, in press).

The way emotions may mix

Picard evokes the *generative mechanism* as the key factor for distinguishing between emotions that may *coexist* (by mixing according to a 'tub of water' metaphor) and emotions that rapidly *switch* from each other in time (by mixing according to the 'microwave oven' metaphor): *"to the extent that two emotions have non overlapping generative mechanisms,.. they can coexist in time"* . Co-existence might be due, as well, to differences in the time of decay

[1] Information Societies Technology Programme IST-1999-29078

between emotions generated by the same mechanism: for instance, 'primary' emotions, like *fear,* and cognitively-generated ones, like *anticipation* (Picard, 2000).

To represent the two ways emotions may mix up, modeling formalisms should therefore enable representing their generative mechanism, the intensity with which they are triggered and the decay of this intensity with time. We claim that this representation power may be achieved with **dynamic belief networks (DBNs)**. We briefly describe in the APPENDIX what are DBNs and how they may be employed to represent emotion triggering, and focus our analysis on *the differences in the generative mechanism of emotions* that may be represented with DBNs; our final goal is to establish a correspondence between the differences that might occur in these mechanisms and the two mixing metaphors proposed by Picard.

Goal-Based Emotion Categories

'Fortune-of-others' emotions (*sorry-for, happy-for, envy and gloating*) may be represented as points in the two-dimensional space ('desirability of the event', 'empathy of the reaction'): happy-for and envy apply to 'desirable' events while 'sorry-for' and 'gloating' apply to negative ones; happy-for and sorry-for are driven by an empathic attitude, while gloating and envy are driven by a contrasting one.

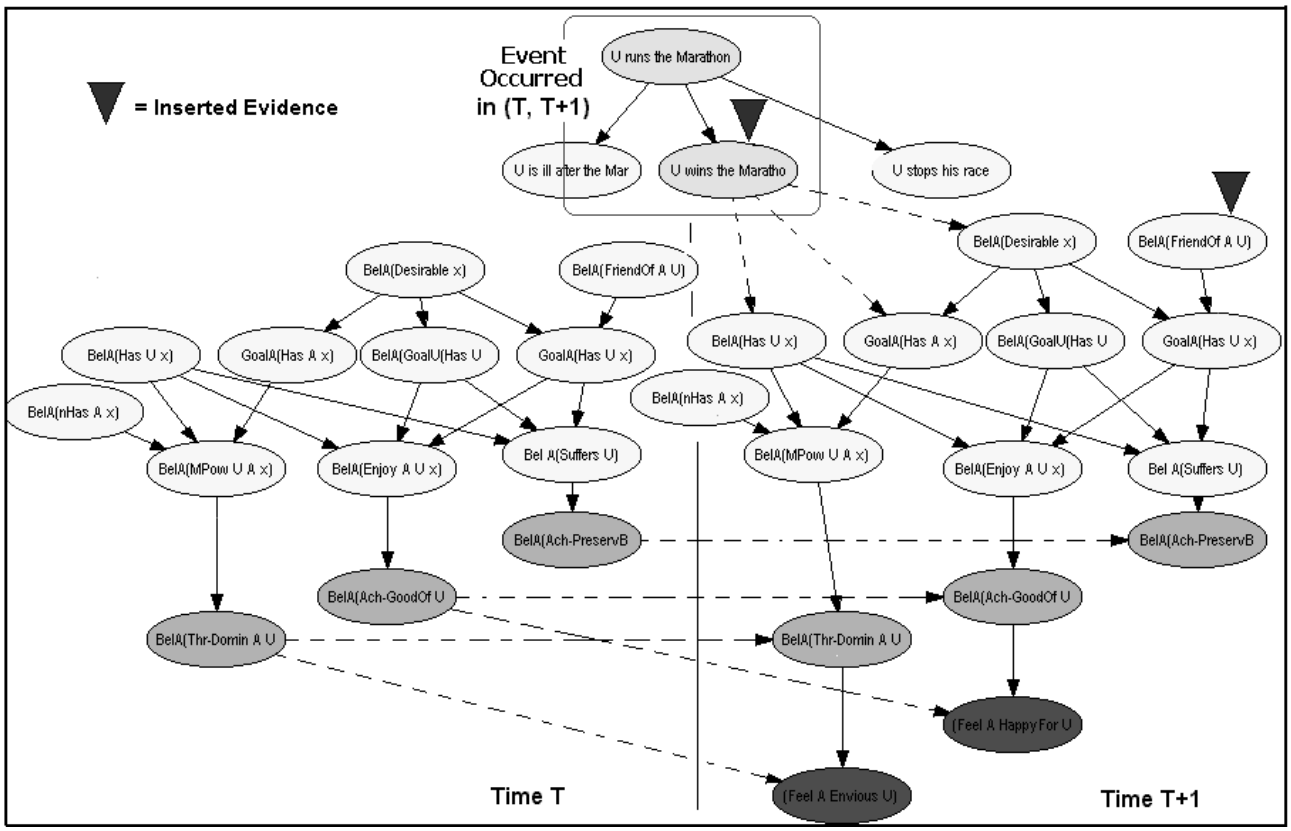

Figure 1: a dynamic belief network for triggering of envy and happy-for

Figure 1 shows the dynamic belief network that models how *happy-for* and *envy* may be triggered in an Agent A who assists to her husband U running the Marathon. This model shows that *happy-for* is triggered by an increasing probability in the belief that the high-level goal of 'desiring the good of others' will be achieved (Ach-GoodOf U). The intensity of this emotion depends on the variation of this probability, that is produced when evidence about some desirable occurred event (the runner comes to the end of the race) is propagated. It depends, as well, on the weight the Agent attaches to achieving that goal; this weight is, in its turn, a function of the Agent's personality (high for *altruistic*

people, low for *egoistic* ones). The emotion of *sorry for* is triggered by the same nodes as *happy-for*, when these nodes have opposite values because the occurring event is *undesirable* (for instance, the runner has a suffering aspect): the high-level goal involved in this case is that of 'preserving others from bad' (Ach-PreserveBad U). *Gloating* is related to envy by a similar triggering mechanism: the high-level goal involved is, in this case, that of 'desiring the bad or others' and 'being preserved from bad', at the same time.

Which 'fortune of others' emotions may mix up, and how? According to our model, *happy-for* and *envy* are examples of potentially 'coexisting' emotions: their generative mechanisms are different in that *the nodes that trigger them*, in the BN, *belong to two nonoverlapping sets*. For instance, the node Goal A (Has A x) does not affect achievement of the goal 'desiring the good of others'. So, Agents who are moderately altruistic and moderately dominant may be moderately envious and moderately happy-for at the same time, when they come to know that a desirable event occurred to a friend. For similar reasons, *sorry-for* and *gloating* may coexist, in particular circumstances. Does the same mixing metaphor apply to other emotion combinations, in this set? Probably not. The generative mechanisms of *happy-for* and *sorry-for*, for instance, are the same: as we said, they are triggered by a different value of desirability of the occurring event and, therefore, by *different values of common nodes* in the BN; therefore, they cannot coexist. However, it might happen that the same event *activates*, in two different time instants, *different kinds of beliefs*. Persons may see positive and negative consequences of the same event, at different times ("*he is going fast! Oh, but he seems to be very tired!*"): they will, in this case, switch more or less rapidly between *happy-for* and *sorry-for* emotions, according to how rapid is the activation of new beliefs about the consequences of the event.

More or less fast switching in time may occur from **Prospect-based** emotions (*fear, hope*) to **Well being** (*distress, joy*) or **Confirmation** emotions (*disappointment, relief*). Belief about achievement or threatening of high-level goals ('Desiring the Good of Self' or 'Preserving Self from Bad') are involved in these cases. Switching from one emotion to another is due to a change in the probability of the belief that a (desirable or undesirable) event will occur, is occurring or occurred; this change may be due to observation of different consequences of this event at different times. In the Marathon example, switching from **fear** to **joy** to **relief** is closely related to the probability with which the runner believes she will win the race; this probability changes from time to time, according to the progress of the race. A 'microwave oven' metaphor may therefore be applied to mixing of emotions originating from the same belief observed at different time instants, in the three categories. The same happens for emotions within the Prospect-based, the Well-being or the Confirmation category, which cannot coexist but between which the Agent may switch. For instance, the Marathon runner was, in different time instants, **hopeful** to win the race and **fearful** about its possible consequences on her health.

According to similar considerations, one may conjecture about mixing of emotions belonging to different categories: for instance, **happy** and **happy-for** may coexist, as well as **sorry-for** and **relief**, … and so on.

Power of Dynamic Belief Networks in Representing the Mixing of Emotions

DBNs enable representing the three kinds of situations that produce mixing of emotions:

a. *concurrent triggering* of emotions occurs when an event has *different consequences* whose beliefs co-occur in the Agent's mind and when these beliefs influence, in their turn, the achievement of different but *compatible* goals[2];

b. *rapid switching* among different emotions occurs when the probability that the Agent attaches to a given belief changes with time or when new beliefs are activated in different time instants: these beliefs may be consequences of the same or of different events.

[2] Compatibility of goals is defined according to psychological plausibility considerations: for instance, it is infrequent to attach a *very high* importance to 'dominating others' and to 'desiring their good', at the same time. Medium weights to both goals might be, on the contrary, plausible.

As our DBNs are cognitive models of the way *all* types of emotions are elicited, we do not make a distinction between primary and cognitively-generated emotions. However, we represent time decay of goals by introducing a link between two nodes representing the same goal at different time instants; the strength of this link may be varied, to simulate various forms of time decay. This enables us to formalise, as well, the case in *which coexistence is due to the fact that, after the same stimulus, emotions tend to take more or less time to disappear* and therefore may coexist for part of the time.

Final remarks

The main difference between our models and other modeling approaches (Elliott and Siegle, 1993 and Gratch and Marsella, 2001, to cite only a few of them) is in the 'grain size' of knowledge represented and in the theory applied to measure the emotion intensities. Rather than representing directly the effect of events on the Agent's affective state, by an application of OCC theory (Ortony et al, 1988), we build a fine-grained cognitive structure, in which the appraisal of what happens in the environment is represented in terms of its effects on the Agent's system of beliefs and goals. The main advantages of this formalism are in the possibility of representing the following aspects of emotion elicitation:

- *temperament and personality:* how these factors affect the Agent's propensity to feel and to show emotions, by influencing the weights it assigns to achieving its goals;
- *social context:* how the Agent's relationship with the context influences emotion triggering by influencing its beliefs and goals;
- *dynamics of the Agent's state:* how the Agent's affective state changes over time, either as a consequence of 'endogenous' or of 'exogenous' factors;
- *response decay*: how emotion felt by the Agent decays, in absence of any new specific stimulus;
- *goal revision*: how triggering of an emotion induces the Agent to revise its goals.

In addition, we demonstrated, in this short contribution, the ability of DBNs in representing the way emotions mix: this is the main issue we would like to discuss with participants to the Symposium.

Main references

B De Carolis, C Pelachaud, F de Rosis and V Carofiglio: Interactive Information Presentation by an Embodied Animated Agent. *Submitted.*

C. Elliott and G. Siegle: Variables influencing the intensity of simulated affective states. *Proceedings of the AAAI Spring Symposium on Mental States '93*, 58-67, 1993.

J. Gratch and S. Marsella: Tears and Fears: Modeling emotions and emotional behaviors in synthetic agents. *Proceedings of the 5th International Conference on Autonomous Agents*, Montreal, Canada, June 2001.

K. Oatley and P.N. Johnson-Laird: Towards a Cognitive Theory of Emotions. *Cognition and Emotion,* 29-50, 1987.

A. Ortony, G.L. Clore and A. Collins: *The cognitive structure of emotions*. Cambridge University Press, 1988.

A Ortony: On making believable emotional agents believable. In R. Trappl and P. Petta (Eds): *Emotions in humans and artifacts*. MIT Press, Cambridge, MA, in press.

J Pearl: *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufman Publishers, 2000.

R Picard: *Affective Computing*. The MIT Press, 2000.

APPENDIX: Outline of the model of emotion triggering, with DBNs.

Emotions are triggered by the belief that a particular important goal may be achieved or threatened. So, the main focus of our simulation is the *change* in the belief about the achievement (or threatening) of goals of an Agent A, over time. In our monitoring system, the cognitive state of A is modeled at the time instants {T, T+1, T+2, …}. Events occurred during the time interval (T,T+1) are observed, to construct a probabilistic model of the new state and reason about emotions that might be triggered by these events. An appropriate formalism for representing dynamic phenomena in conditions of uncertainty is that of *dynamic belief networks* (DBNs). DBNs are based on the idea that time is divided into time slices, each representing the state of the modelled world at a particular instant or interval; this state is described by means of a static belief network. Two consecutive time slices are linked by arrows between the corresponding domain variables we want to monitor. When something changes in the world, the network is extended for an additional time slice (Pearl, 2000). As a consequence, its structure may change over time in a number of ways.

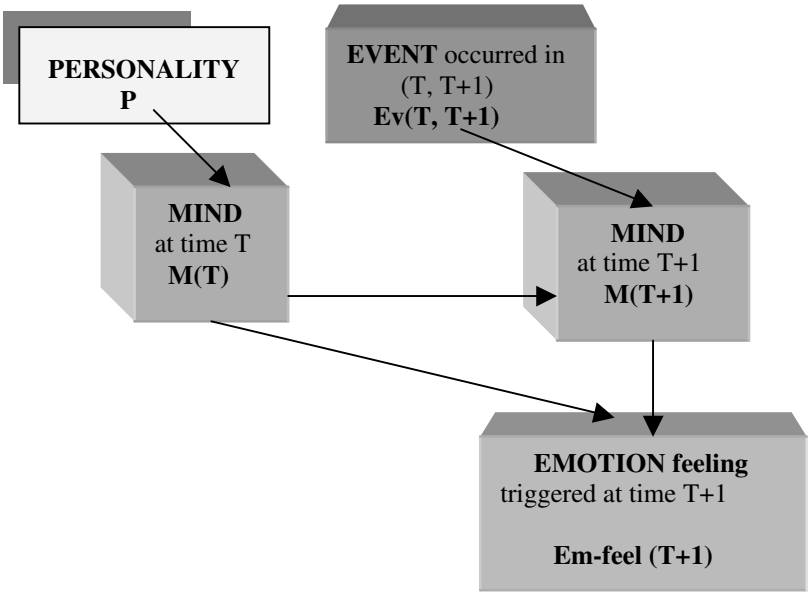

Figure 2: General Structure for the Emotion Monitoring System

Figure 2 shows the general structure of our model, that includes the following static components:

- M(T) represents A's Mind at time T, with its beliefs about the world and its goals;
- Ev(T,T+1) represents the event occurred in the time interval (T, T+1), with its causes and consequences;
- M(T+1) represents A's Mind at time T+1;
- Em-feel(T+1) represents the fact that a particular emotion is activated, in A, at time T+1.

M(T+1) depends on M(T) and the event occurred in the interval (T,T+1). The feeling of emotions depends on both M(T) and M(T+1). We calculate the intensity of emotions as a function of two parameters: (i) the *uncertainty* in A's beliefs about the world and, in particular, about the possibility that some important goal is achieved or threatened, and (ii) the *utility* assigned to achievement of this goal. More in depth, if:

- Bel A Ach(Gi) is A's belief that the goal Gi will be achieved;
- P(Bel A Ach(Gi)) and P*(Bel A Ach(Gi)) are the probabilities that A attaches to this belief, respectively before and after the event Ev occurred;
- W(Ach(Gi)) is the weight that A attaches to achieving Gi,

then, according to the utility theory (Pearl, 1990), the *variation of intensity in the emotion (ΔIe)* may be calculated as:

$$\Delta Ie = [P(Bel\ A\ Ach(Gi)) - P^*(Bel\ A\ Ach(Gi))] * W(Ach(Gi))$$

In other words, ΔIe is the product of the change in the probability that Gi will be achieved, times the weight of the goal.