



## Model for Artificial Emotion from a Utilitarian Point of View

Utsav Datta

*Amity school of engineering and technology, Amity university Kolkata*

*\*Corresponding Author's Email: [utsavdatta07@gmail.com](mailto:utsavdatta07@gmail.com)*

### Abstract

Implementation of Emotion in Artificial intelligence is an open field and intensive research is being done on this topic for the past decades. This paper proposes a model for triggering six basic emotions described by Paul Ekman by using the concept of Pleasure and Pain. Using pleasure-pain, we classify events into Desirable and Undesirable events. The model uses a probability graph for determining the probability of a certain event happening through a certain path and the intensities of the emotions are a function of pleasure-pain and the corresponding probability derived from the graph. In the end of the paper, we conclude through an illustration that the model is capable of capturing the dynamic nature of emotions.

**Keywords:** *Artificial emotion problem, emotion model, emotion simulation and excitation, emotional agents.*

### 1. Introduction

In computer science and Engineering, Artificial intelligence is an emerging field of prime importance. Apart of making a system more computationally intelligent, research is going on to implement emotions in them. Emotions are extremely important in human beings. A study conducted by Damasio, concluded that humans who lack emotion cannot differentiate between good and bad and cannot take small decisions [1]. This paper [2] concluded from their study that, high EI robots and humans received a higher rating of being trustworthy from humans than robots and humans with low EI. There are several definitions available for emotions, Oxford Dictionary defines emotion as “A Strong feeling deriving from one’s circumstances, mood, or relationship with others.” Cabanac thought of emotions as a mental experience [3] and Lazarus told emotions are based on person-environment relationship [4].

While there are many theories of Emotions, almost everyone agrees that Cognition is an important part of emotion. A system needs to understand its surrounding and be “aware” of its environment. The events that are happening in the environment are responsible for triggering emotions in a system. This led to “Appraisal theory” which suggests that emotions are triggered when particular appraisals are made [5]. That is, the theory suggests that emotions are only triggered when the system assesses its environment for certain conditions. Many psychological models were developed based on the appraisal theory, such as the OCC model [6], Roseman’s theory of appraisal [7], Scherer’s multi-level model [5] etc. and based on such psychological models based on appraisal theory, many computational

emotional models have been proposed by researchers to simulate emotions in artificial systems. Some of the existing models based on Appraisal theory are the FLAME model, EMA model and the EmET model. We would discuss them in a very brief manner.

FLAME, which is the acronym for Fuzzy Logic Adaptive Model of Emotion is a computational model of emotion based on the appraisal theory. The model has 3 components, which are “Emotional component”, “learning component” and the “decision-making component”. The model uses fuzzy logic to generate emotional intensities by assessing the impact of the events on a goal [8].

EMA model, which stands for Emotion and Adaptation, is a computational model that uses appraisal variables. The combination of these variables give rise to emotions and their intensities. The appraisal variables defined in EMA are: Perspective, Desirability, Likelihood, Casual attribution, Temporal Status, Controllability and Changeability [9]. The emotions defined in EMA are: Joy, Hope, Fear, Dismay and Anger.

EmET model, which stands for Emotion Elicitation and Emotion Transition Model, is also based on appraisal theory and it targets 5 primary emotions. i.e. Happy, Sad, Fear, Surprise and Anger. This model uses linguistic variables and IF-THEN rule to generate emotions and their intensities. The intensities generated are also in linguistic terms [10].

The model that is being proposed through this paper is inspired from the EMA model by Gratch and Marsella. The EMA model uses Desirability and Likelihood as Appraisal variables to compute emotions. We want to provide a formal model which uses Desirability and Likelihood to trigger emotions. We use the concept of Utilitarianism for computing desirability and a traversable probabilistic graph for computing Likelihood. We consider the 6 basic emotions that were defined by Paul Ekman for our model [11].

## 2. Proposed model

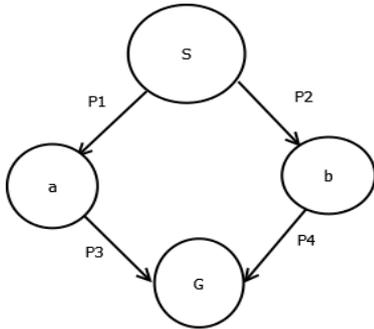
Jeremy Bentham put forward a philosophical theory called utilitarianism, which was based on the idea that, actions are justified when they lead to pleasure [12]. This concept can be implemented for events as well. When confronted with multiple events, the event which gives the max pleasure is the one that is desired. The greatest happiness principle, which can be thought of as the heart of utilitarianism states that right actions are actions which promote happiness (amount of pleasure increases while absence of pain) and wrong actions are actions that decrease happiness (increase pain while there is absence of pleasure) [13]. For our model, we are considering that every event has a corresponding net pleasure-pain value that would be felt by our system if it experiences that event. The new thing that we are introducing is the concept of a cluster of events. We know that events are not isolated in nature. One event leads to another. If we consider a GOAL event for our system, then that GOAL can be reached through different events but obviously the net amount of pleasure-pain and also the probability of successfully reaching the GOAL would be different for different paths. We define the net pleasure-pain for any event as:  $x = \sum \text{pleasure} - \sum \text{pain}$ .

This  $\sum \text{pleasure}$  and  $\sum \text{pain}$  can be found by an algorithm known as Felicific Calculus [14].

For creating our cluster, 1<sup>st</sup> the system has to consider a GOAL event. This GOAL event can be DESIRED or UNDESIRED. It would be DESIRED for the value of  $x > 0$  and UNDESIRED for the value of

$x < 0$ . From a real-world perspective, a desired goal would be passing a test while similarly an undesired goal would be failing a test.

We would consider a single starting point whose  $x = 0$ , and the probability of moving to any of the branched events from the starting point is 1. Between the starting point and GOAL, there are intermediate nodes representing the events that needs to be faced before we can reach to the GOAL. If we represent the cluster in a form of a graph, then each node representing an event will have a net pleasure-pain value and each edge connecting 2 nodes would have a probability value.



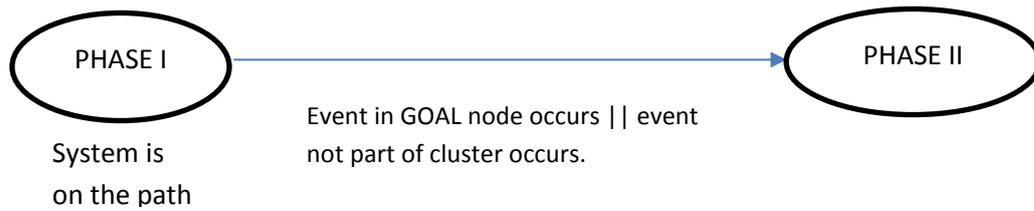
**Fig. 1. Example 1**

In the adjacent figure, we see that the system can reach GOAL “G” from “S”, through different events “a” and “b”. Suppose that “a” causes a net pleasure-pain of -A and “b” causes a net pleasure-pain of -B, then we can say that “G” can be reached through “a” with a probability of  $P1 \cdot P3$  and a cost of -A, while “G” can be reached through “b” with a probability of  $P2 \cdot P4$  and a cost of -B.

Which path to choose from in a cluster can be depended on various factors such as the current emotional state, the maximum net pain a system can take etc. however, from a utilitarian view we can say that *the system will endure temporary pain in the short run for a high probability of achieving high pleasure in the long run, or, the system will endure pain in the short run to avoid higher level of pain in the long run*. Just like a person who has an exam coming would ideally decide to study even though studying would cause pain, when compared to enjoying but through studying the person achieves a higher probability of passing which would bring more pleasure in the long run, compared to failing, which would bring pain.

The path between a Starting node and GOAL can be arbitrarily long. That means it can contain many numbers of intermediate nodes. The cost of the entire path is the summation of net pleasure-pain of the intermediate nodes in that path. We denote the cost of the  $i^{th}$  path by  $\beta_i$ , and the probability of reaching the GOAL through  $i^{th}$  path by  $\psi_i$ . Also, the intermediate nodes can be of 2 types, either it can represent events that are selectable by the system or it can represent an event that is out of hand of our system i.e. the system cannot choose the node, it can only move through it if it occurs. The nodes that are not selectable are identified with an extra variable, say  $\lambda$ . We use  $x, \beta_i, \psi_i$  and  $\lambda$  to define the 6 emotions.

Before defining the emotions, we would define 2 phases. *PHASE I* is the phase when the system has not reached the GOAL node and is still inside the cluster, on the path for reaching the GOAL. *PHASE II*, however, is the phase which is activated after the event in GOAL node has actually occurred, or an event that is not part of the cluster has occurred, thus breaking the path.



Paul Ekman defined 6 basic emotions, which are: HAPPINESS, SADNESS, SURPRISE, FEAR, ANGER and DISGUST. Phase I, contains the emotions HAPPINESS, SADNESS and FEAR. Phase II, contains the emotions HAPPINESS, SADNESS, SURPRISE, ANGER and DISGUST.

The emotions defined in Phase I are dynamic in nature, as in, the values keep on changing as the system moves along the path.

**Phase I**

if we consider any path  $[S \rightarrow E_1 \rightarrow E_2 \rightarrow \dots E_n \rightarrow G]$ , then for computing the emotions, firstly the net pleasure-pain of the GOAL event must be known and must be classified as DESIRABLE or UNDESIRABLE.

IF  $x > 0$  THEN desirable

ELSE undesirable

The probability  $\psi_i$  and the path-cost  $\beta_i$  are time dependent. They change values as the system progresses on the path. If at starting node "S", we consider a set M which contains the probabilities of the path as its elements and a set N, which contains the probabilities of the edges that the system has already passed through in the path, then at any time  $t$ , the probability is

$$\psi_{i(t)} = \frac{\prod_{p \in M} p}{\prod_{p \in N} p} \quad \text{where } p \text{ is denoting probability.}$$

Likewise, the path-cost at any time  $t$  can be computed by summation of all the individual node-cost (net pleasure-pain) of the events (nodes) that the system has gone through. If  $\beta_i$  is the net pleasure-pain of a particular node, then  $\sum_{i=1}^k \beta_i$  is the total cost after traversing  $k$  number of nodes.

We now move on to defining the emotions for Phase I. we start with HAPPINESS. From a utilitarian view, happiness is the presence of pleasure and absence of pain. We define happiness for both DESIRED and UNDESIRED GOAL events.

HAPPINESS FOR DESIRED GOAL:

$$H = \begin{cases} 0 & \text{if } 0 \leq \psi_{i(t)} < 0.4 \\ \frac{|\psi_{i(t)} * (x + \beta_{i(t)})|}{2} & \text{if } 0.4 \leq \psi_{i(t)} \leq 0.6 \\ |\psi_{i(t)} * (x + \beta_{i(t)})| & \text{if } 0.6 < \psi_{i(t)} \leq 1 \end{cases}$$

HAPPINESS FOR UNDESIRED GOAL:

$$H = \begin{cases} 0 & \text{if } 0.6 < \psi_{i(t)} \leq 1 \\ (1 - \psi_{i(t)}) * \frac{|(|x| + \beta_{i(t)})|}{2} & \text{if } 0.4 \leq \psi_{i(t)} \leq 0.6 \\ (1 - \psi_{i(t)}) * |(|x| + \beta_{i(t)})| & \text{if } 0 \leq \psi_{i(t)} < 0.4 \end{cases}$$

Similarly, SADNESS is the presence of pain and absence of pleasure.

SADNESS FOR DESIRED GOAL:

$$S = \begin{cases} 0 & \text{if } 0.6 < \psi_{i(t)} \leq 1 \\ (1 - \psi_{i(t)}) * \frac{-\beta_{i(t)}}{2} & \text{if } 0.4 \leq \psi_{i(t)} \leq 0.6 \\ (1 - \psi_{i(t)}) * (-\beta_{i(t)}) & \text{if } 0 \leq \psi_{i(t)} < 0.4 \end{cases}$$

SADNESS FOR UNDESIRED GOAL:

$$S = \begin{cases} 0 & \text{if } 0 \leq \psi_{i(t)} < 0.4 \\ \frac{|\psi_{i(t)} * (x + \beta_{i(t)})|}{2} & \text{if } 0.4 \leq \psi_{i(t)} \leq 0.6 \\ |\psi_{i(t)} * (x + \beta_{i(t)})| & \text{if } 0.6 < \psi_{i(t)} \leq 1 \end{cases}$$

FEAR is a little tricky to define. Most theories agree that fear arises due to uncertainty. We use this and define fear for DESIRED GOAL as: when  $\psi_{i(t)}$  is less than 0.5, then the system would have a fear component which equals to  $(1 - \psi_{i(t)}) * x$ . This is basically the “fear of missing out the pleasure brought on by GOAL event”. For UNDESIRED GOAL, when  $\psi_{i(t)}$  is greater than 0.5, then the system would have a fear component which equals to  $\psi_{i(t)} * |x|$ , which denotes “fear of the pain brought on by the undesired event”.

FEAR FOR DESIRED GOAL:

$$F = \begin{cases} 0 & \text{if } \psi_{i(t)} > 0.5 \\ (1 - \psi_{i(t)}) * x & \text{if } \psi_{i(t)} < 0.5 \end{cases}$$

FEAR FOR UNDESIRED GOAL:

$$F = \begin{cases} 0 & \text{if } \psi_{i(t)} < 0.5 \\ \psi_{i(t)} * |x| & \text{if } \psi_{i(t)} > 0.5 \end{cases}$$

The emotion FEAR, only stays in the system as long as the system is in phase I. FEAR becomes zero when the phase is transitioned to Phase II.

We have defined the phase I emotions for both DESIRED and UNDESIRED GOALS. The values computed for the emotions for phase I, only stays in the system as long as the system is in phase I. Once the values of emotions are computed for phase II, the phase II values overwrite the phase I values. Emotions such as surprise, anger and disgust are not “felt” by the system while it is in phase I, i.e. while the system is in path to reach the GOAL.

Suppose, at time  $T_2$ , the system either reaches the GOAL node or moves out of the cluster (event outside the cluster occurs, breaking the path). Then at time  $T_2 + \delta t$  the PHASE II is initiated.  $\delta t$  is the reaction time and is very small. The new intensities for the PHASE II emotions are computed according to the following equations and the new intensities replaces the PHASE I intensities.

## PHASE II

In phase II, we first consider a Boolean  $B_o$ , the value of the Boolean is set to 0 if the system moves out of cluster and the value is set to 1 if the event in the GOAL node, whether DESIRED or UNDESIRED occurs.

Surprise: considering that surprise is a state of excitation that the system goes in when there was a high probability of something happening, but that event does not happen. For computing surprise, we first need the current  $\psi_{i(t)}$  value. The  $\psi_{i(t)}$  value gives us the probability of the GOAL node happening at that instance of time. We compare it with the following condition to find the SURPRISE state.

SURPRISE:

$$\text{If } \psi_{i(t)} \geq 0.8 \ \&\& \ B_o = 0 \\ \text{SURPRISE} = 1$$

$$\text{If } \psi_{i(t)} \leq 0.2 \ \&\& \ B_o = 1 \\ \text{SURPRISE} = 1$$

Anger: Many theorists such as Ortony et al. [6] proposed that anger is a complex emotion. In this paper we are considering anger as a combination of SADNESS and  $\lambda$ .  $\lambda$  is a symbol that denotes that a “non-selectable” node has occurred on the path. A path can have ‘ $n$ ’ number of non-selectable nodes and therefore anger would become SADNESS +  $n(\lambda)$ . The value of  $\lambda$  is dependent on sadness.

$$\text{Anger} = S + \lambda \\ \lambda = f(S) \text{ where } S \text{ is the intensity of sadness.}$$

However, the intensity of ANGER is 0, when there are no “non-selectable” nodes present in the path. The absence of such nodes on the path means that the system itself has chosen which events to go through. A real-life example of such a case would be that, while a student can choose to study or not but cannot choose whether the exam paper would be hard or easy. The event that represents the paper is ‘hard’ or ‘easy’ is a  $\lambda$  event.

Disgust: Disgust is an emotion that provides repulsion from a certain condition. It is generally considered that disgust is an evolved emotion to reject certain foods that are harmful [15]. Because of this repulsive nature, in this model, DISGUST has a direct effect on the cost of a path. The basic ideology is that, if a path repeatedly gives failing results then overtime the system should understand not to take that particular path. Considering that the probability is independent of the system then we achieve this rejecting property by increasing the cost of the path (increasing the net pain of the path). Note- for DESIRED GOAL, the failing result is not reaching the GOAL node, while for UNDESIRED GOAL, the failing result is reaching the GOAL node.

The mathematical description for DISGUST is as follows: let  $D_i$  be a variable storing the value of DISGUST for the  $i^{th}$  path. Every time a system goes through that particular path and failing result occurs, the value of  $D_i$  is incremented by a certain factor  $\mu$ .  $\mu$  is a function of SADNESS which is caused due to failing result.

$$D_i = D_i + \mu \text{ where } \mu = f(S)$$

Now, if we consider the number of nodes in a path to be ' $n$ ', then after failing, the net pleasure-pain value of the individual nodes in the  $i^{th}$  path is updated as

$$\beta = \beta - \frac{D_i}{n} \text{ where } \beta \text{ is the net pleasure - pain of a particular node}$$

If the system continuously achieves a failing result from a particular path, then the  $\beta$  value of the nodes residing in the path become more and more negative with each failing. This denotes that the nodes deliver more and more pain to the system. This process iterates until the nodes become so negative that the system rejects the said nodes and tries a new path which causes less amount of "pain".

Happiness: While defining HAPPINESS in PHASE I, a  $\psi_{i(t)}$  component was present which took into account the probability of reaching a GOAL node. However, in phase II we don't require a probability component because in this phase the system has either reached the GOAL node or has moved away from the path reaching the GOAL node.

For DESIRED GOAL

$$H = \begin{cases} 0 & \text{if } B_o = 0 \\ |x + \beta_{i(t)}| & \text{if } B_o = 1 \end{cases}$$

For UNDESIRED GOAL

$$H = \begin{cases} 0 & \text{if } B_o = 1 \\ (|x| + \beta_{i(t)}) & \text{if } B_o = 0 \end{cases}$$

Sadness: the intensity of SADNESS is computed using the same concept as that of HAPPINESS. We omit the probability component in phase II.

For DESIRED GOAL

$$S = \begin{cases} 0 & \text{if } B_o = 1 \\ -\beta_{i(t)} & \text{if } B_o = 0 \end{cases}$$

FOR UNDESIRED GOAL

$$S = \begin{cases} 0 & \text{if } B_o = 0 \\ |x + \beta_{i(t)}| & \text{if } B_o = 1 \end{cases}$$

Therefore, we have provided the mathematical formulation of the emotions above for the 2 phases. We have considered that the computed intensities of the emotions must be a positive value. i.e. The value of intensity lies between  $[0, \infty)$ . If the intensity for an emotion comes out to be less than 0, then we consider the intensity to be 0. For instance, when we compute SADNESS for a DESIRED GOAL, if the system has experienced "pain" which is negative in nature, only then the intensity value of SADNESS comes out to be positive due to the negative sign in  $-\beta_{i(t)}$ . If the system has experienced a net pleasure in its path, then the intensity would come out to be negative, hence the intensity of SADNESS would be 0. This logic is correct from the utilitarian perspective which suggests that sadness comes from pain, hence if the system gains pleasure from the path, intensity of SADNESS would be 0.

### 3. Decay of Emotions

The intensities of the emotions keep changing with the change in pleasure-pain and probability of reaching the GOAL. However, if the system’s emotional intensities do not change over a certain period of time it is natural that the said intensities would decrease gradually. The emotion FEAR only remains in the system till it is in phase I. It becomes 0 in phase II. The emotion SURPRISE has a discrete value of 1 or 0. The decay of other emotions can be formulated to be linear with time.

$$H = H - H_f(t)$$

$$S = S - S_f(t)$$

$$A = A - A_f(t)$$

$H_f, S_f, A_f$  are variables & represent the deexcitation factors for HAPPINESS, SADNESS and ANGER respectively. These are basically small integer values that determine how rapidly the emotions decay in the system w.r.t time "t".

The intensity of DISGUST is computed for a particular path. As stated before, this disgust increases the “pain” value for a path which leads to repeated failure for achieving the rejecting property. However, if the path provides a successful result, then the intensity of Disgust for the path must decrease. The intensity of DISGUST increases according to the equation:  $D_i = D_i + \mu$  where  $\mu = f(S)$ . Similarly, on achieving success, the intensity decreases as  $D_i = D_i - \mu'$  where  $\mu'$  is a function of happiness achieved by traversing through that particular path. The new net pleasure-pain values for the individual nodes in the path are updated as before by increasing the net pleasure-pain by a value proportional to  $\mu'$ .

### 4. Example and Discussions

Above, we provided a formal model for excitation and deexcitation of emotions in a system. In this section of the paper we are providing an example that shows the transition of emotions in the system when faced with a real-world situation.

The situation consists of our system reaching university on time. Therefore, our GOAL event (G) is “reaching university on time”. It is a DESIRED GOAL therefore it must provide pleasure. To illustrate our example, we are using arbitrary pleasure and pain values. We are assuming that the node (G) will provide a pleasure of +120. There are a number of intermediate nodes between (S) and (G).

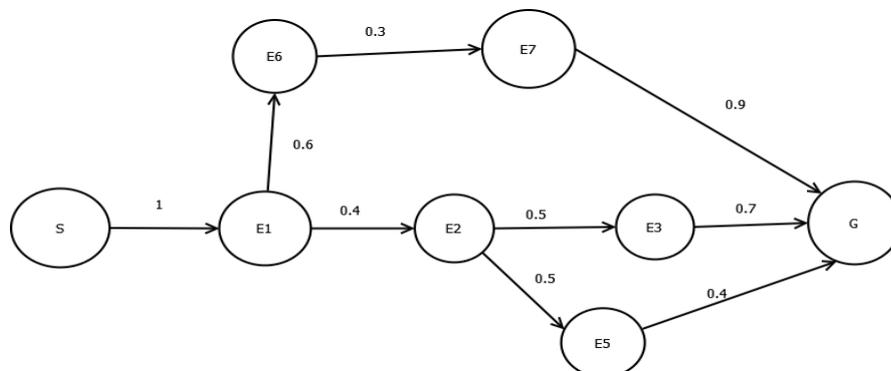


Fig. 2. Example 2

These intermediate nodes represent events. Node (E1) represents “going outside”, Node (E2) represents “reaching bus stop on time”, Node (E6) represents “reaching train station on time”, Node

(E7) represents “taking train”. The node (E3) and (E5) are both  $\lambda$  events representing “taking bus 1” and “taking bus 2” respectively.

Each of the nodes has a net pleasure-pain value ( $\beta$ ) assigned to them which are:

**Table 1.  $\beta$  values for nodes in Example 2**

NODE	$\beta$
(S)	0
(E1)	-20
(E2)	-60
(E3)	0 ( $\lambda$ )
(E5)	0 ( $\lambda$ )
(E6)	-70
(E7)	-80

THE system moves to (E1). At this node, branching takes place and the system has to choose whether it wants to “reach the train station” or “reach the bus stop on time”. Here the graph search algorithm takes over. The system knows path  $\{E1 \rightarrow E6 \rightarrow E7\}$  has total  $\beta$  of -170,  $\{E1 \rightarrow E2 \rightarrow E3\}$  has total  $\beta$  of -80 and path  $\{E1 \rightarrow E2 \rightarrow E5\}$  has total  $\beta$  of -80. The system gains a pleasure of +120 from (G) hence, from a utilitarian view it is not “worth it” for the system to move to (E6) as path through (E6) has higher “pain” (-170) than pleasure gained from Goal Node(G) i.e. +120. Due to this, the path  $\{E1 \rightarrow E6 \rightarrow E7\}$  gets deleted from the cluster.

$t = E1 \rightarrow$  At (E1) the  $\beta_{i(t)}$  and  $\psi_{i(t)}$  are: 0 and 0.22 respectively.

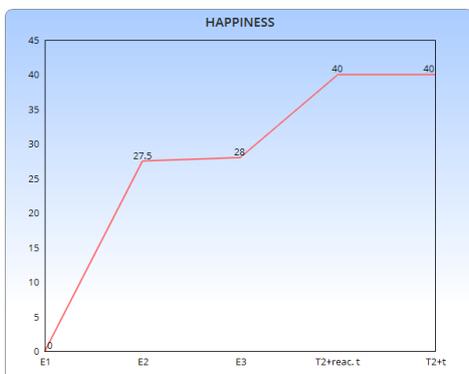
$t = E2 \rightarrow$  At (E2) the  $\beta_{i(t)}$  and  $\psi_{i(t)}$  are: -20 and 0.55 respectively.

At node (E2) another branching occurs, but this one is a  $\lambda$  event. Unlike the previous branching the system cannot choose between them this time. We are assuming that event represented by (E3) occurs.

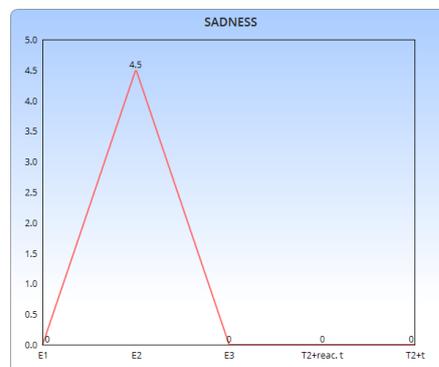
$t = E3 \rightarrow$  At (E3) the  $\beta_{i(t)}$  and  $\psi_{i(t)}$  are: -80 and 0.7 respectively.

$t = T_2 \rightarrow$  At (G) the  $\beta_{i(t)}$  is -80. There is no  $\psi_{i(t)}$  as the system has reached the goal successfully.

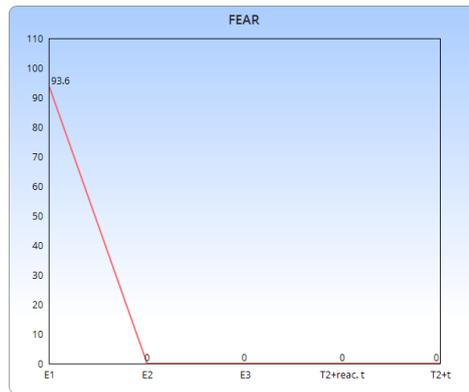
In our illustration we have considered that the system reaches the DESIRED GOAL successfully. We get an intensity of the emotions in each of the nodes and below we have provided a graph representing the level of the emotions in each of the stages. The NODES are plotted on x axis and the intensities are plotted on y axis.



**Fig. 3. Happiness for success case**



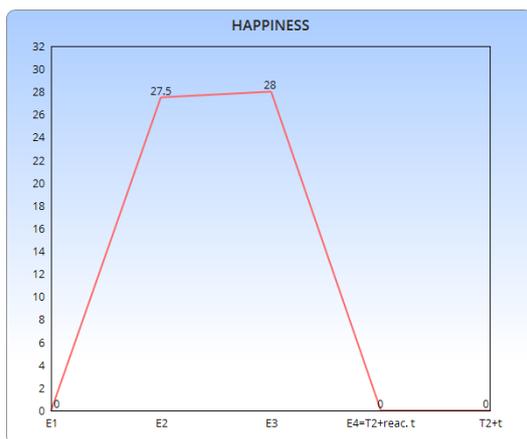
**Fig. 4. Sadness for success case**



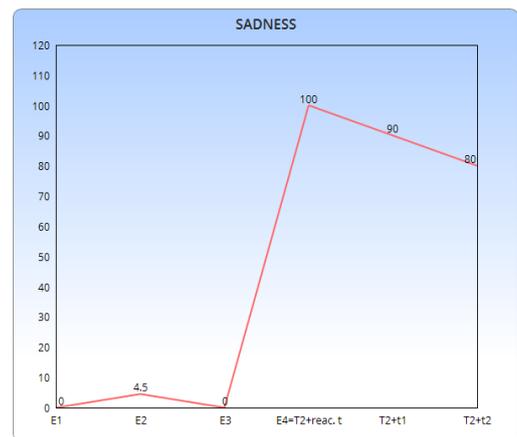
**Fig. 5. Fear for success case**

The system achieves a successful result; therefore, the intensities of ANGER and DISGUST are 0. However, if we consider that another NODE (E4), which denotes the event “Bus 1 breaking down” happens after NODE (E3) then the system encounters a failing result through the path. The phase is immediately transitioned to PHASE II from PHASE I, because (E4) was not a part of the cluster. In this case, we would get different values of emotions.

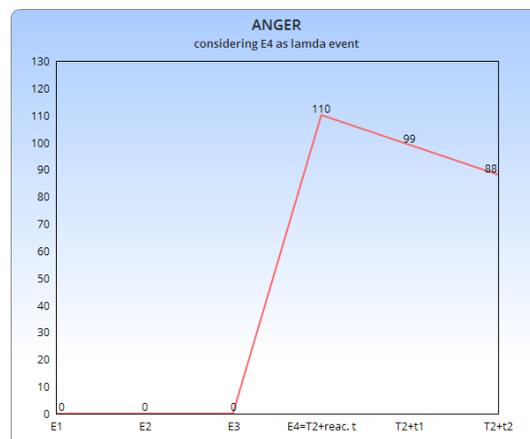
We plot the graphs for such a case where failing result is achieved through the path.



**Fig. 6. Happiness for failed case**



**Fig. 7. Sadness for failed case**



**Fig. 8. Anger for failed case**

1. we have considered that (E4) has a pain value of  $-20$  and  $\lambda = \frac{S}{20}$ . Since in the path, 2 lambda events were present hence, anger at  $T_2 + \delta t$  becomes  $Anger = 100 + 2(\frac{100}{20})$
2. The graph of FEAR remains the same as it was in the previous case.
3. For disgust if we consider  $\mu = \frac{S}{5}$  and the present  $D_i$  value for this path as 0, then after this failing the net pleasure-pain value for each of the nodes in this path is decremented by 20.

Then because of DISGUST, the values of the NODES become:

**Table 2. Updated  $\beta$  values for nodes**

NODES	$\beta$
E1	$-20-20= -40$
E2	$-60-20= -80$
E3	$0-20= -20$

Now, if we again iterate our illustration, then we can see that if the system is given choice to choose its path, then the path  $\{E1 \rightarrow E2 \rightarrow E3\}$  will be rejected. Just like the path  $\{E1 \rightarrow E6 \rightarrow E7\}$ , due to the high amount of pain, the path is not “worth it” anymore. However, we have defined the NODE (E3) as a lambda node, hence it can’t be selected or dis-selected. In the next iteration the system would “hope” for (E5), because the total cost of the path  $\{E1 \rightarrow E2 \rightarrow E5\}$  is now less than  $\{E1 \rightarrow E2 \rightarrow E3\}$ . But if the system has to go through  $\{E1 \rightarrow E2 \rightarrow E3\}$  and success occurs, then DISGUST would go down because of HAPPINESS, thus bringing the cost of the path down eventually.

From the illustrated example, we have shown that the model proposed captures the dynamic nature of emotions. Nonetheless, there are some flaws that are present in this model that we would like to address. The first flaw is the fact that in a real-world situation, emotions are rarely influenced by a single GOAL and at an instance of time, a person can have multiple GOALS. There can be situations where the emotions can be influenced by more than 1 cluster of events. The second flaw is the ambiguity caused because of felicific calculus. Although felicific calculus and the concept of pleasure and pain is being used widely in the field of AI [16], but still there is no general rule/procedure available to determine how much pleasure or pain does a system get due to a particular event. This different values of pleasure-pain in different systems will cause them to have different intensities of emotions for a same GOAL and path. The third flaw is the range of the emotions, which lies between  $[0, \infty)$ . If we want to design a set of rules which takes in the current emotional state and perform actions based on them, then it is necessary to bring down the intensities into a particular range. The range will depend on the set of rules using the intensities.

#### 4. Conclusion and Further Research

This paper presents a formal model for computing intensities of 6 emotions which are Happiness, Sadness, fear, Anger, Disgust and Surprise. Based on mainly two appraisal variables, desirability and likelihood we were able to capture the dynamic nature of the emotions. We used utilitarian pain and pleasure to model the desirability and thus the intensities. We also proposed how a flow of events may occur which leads to excitation of emotions through a traversable graph where pain received from a path acts as the path’s cost. There are plenty of practical applications of the theoretical model

presented in this paper such as selection of right actions in an Intelligent system based on emotions, implementing morality in artificial intelligence etc.

Despite the limitations of the model stated in the previous section, the results of our illustration strongly suggest that the model is capable of increasing the intensities of the right emotions for the events. The future studies should aim to replicate the results for a broad spectrum of emotions. Certain psychological models such as the Plutchik's Wheel where the base emotions can be mixed to form new emotions [17], provide fascinating reference points for developing a more complex computational model for simulating emotions in an Artificial System. The paramount importance of emotion in an Intelligent being warrants further investigation in the subject.

## References

- [1] Damasio, A (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*. Grosset/Putnam, New York.
- [2] Fan, L., Scheutz, M., Lohani, M., McCoy, M., & Stokes, C. (2017, August). Do We Need Emotionally Intelligent Artificial Agents? First Results of Human Perceptions of Emotional Intelligence in Humans Compared to Robots. In *International Conference on Intelligent Virtual Agents* (pp. 129-141). Springer, Cham.
- [3] Cabanac, M. (2002). What is emotion?. *Behavioural processes*, 60(2), 69-83.
- [4] Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American psychologist*, 37(9), 1019.
- [5] Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.
- [6] Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge university press.
- [7] Roseman, I. J. (1996). Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cognition & Emotion*, 10(3), 241-278.
- [8] El-Nasr, M. S., Yen, J., & Ioerger, T. R. (2000). Flame—fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-agent systems*, 3(3), 219-257.
- [9] Gratch, J., & Marsella, S. (2005). Evaluating a computational model of emotion. *Autonomous Agents and Multi-Agent Systems*, 11(1), 23-43.
- [10] Jain, S., & Asawa, K. (2015). EmET: emotion elicitation and emotion transition model. In *Information Systems Design and Intelligent Applications* (pp. 209-217). Springer, New Delhi.
- [11] Ekman, P. (1992). Are there basic emotions?.
- [12] Driver, J. (2009). The history of utilitarianism.
- [13] Mill, J. S. (2016). Utilitarianism. In *Seven masterpieces of philosophy* (pp. 337-383). Routledge.
- [14] Mitchell, W. C. (1918). Bentham's felicific calculus. 1968, 30-48.
- [15] Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and social psychology bulletin*, 34(8), 1096-1109.
- [16] Majot, A. M., & Yampolskiy, R. V. (2014, May). AI safety engineering through introduction of self-reference into felicific calculus via artificial pain and pleasure. In *2014 IEEE International Symposium on Ethics in Science, Technology and Engineering* (pp. 1-6). IEEE.
- [17] Chafale, D., & Pimpalkar, A. (2014). Review on developing corpora for sentiment analysis using plutchik's wheel of emotions with fuzzy logic. *International Journal of Computer Sciences and Engineering (IJCSE)*, 2(10), 14-18.