The **Artificial Emotion Engine** ™, **Driving Emotional Behavior.**

Ian Wilson

Artificial-Emotion.Com
http://artificial-emotion.com
ianw@artificial-emotion.com
November 1999.

Abstract.

The Artificial Emotion Engine is designed to generate a variety of rich emotional behaviors for autonomous, believable, interactive characters. These behaviors are in the form of facial and body gestures, motivational actions and emotional internal states. It will help developers populate simulations with emotional characters while freeing them of the need for an in depth knowledge of neurophysiological systems. Technologies, like the Artificial Emotion Engine, are needed to help fuel the growth of the interactive entertainment industry from a niche medium to the mass-market medium it is becoming. Mass-market products will require the use of social interaction and engaging emotional characters to facilitate the creation of some of their most powerful experiences. Many of the technologies currently employed to produce character behavior, finite state machines (FSM's), scripted conditional logic etc are being stretched beyond their limits. Robotic, one-dimensional characters with limited sets of behaviors populate many interactive experiences, this problem needs to be addressed. In this paper we present an overview of the engine and a review of the neuro-physiological model employed by the engine.

Introduction.

Recent advances in neuroscience have begun to pave the way for the simulation of emotion. We define emotion as being comprised of two conceptual, and in some senses physical, properties. These are cognitive emotions and non-cognitive or innate emotions. We deal, primarily, with non-cognitive emotions in the first version of the engine. Dealing with cognition is the 'Holy Grail' of A.I but as yet it is some way off and it is fraught with problems and complications. As yet, neuroscience has no clear understanding of our cognitive systems, symbolic representations, memory and how these work. This makes any attempt at simulation speculation at best. Simulation of non-cognitive emotions, while less expressive than cognitive emotions, is an attainable goal in the near term.

Recent work is showing us that emotion is an integral part of our brains overall functioning. There is no clear delineation between thinking and feeling they are both neural functions that interact and feedback as a unified system, the brain. So what are we looking to achieve with this engine?

Because this system is being developed as an engine it needs to work in a variety of applications, many of which will not previously have been attempted, especially socially based applications. This constrains the design of the engine but also takes some of the burden away from it. The engine provides, from an action perspective, the characters current motivations. The designer or developer can then use this information to determine what the character will do next.

We conceptually divide emotions into three layers of behavior. At the top level are what we term reactions or momentary emotions; these are the behaviors that we display briefly in reaction to events. For example when someone strikes us. The next level down are moods, these are prolonged emotional states caused by the cumulative effect of momentary emotions, that is signals of punishment and reward. Underlying both of these layers and always present is our personality; this is the behavior that we generally display when no momentary emotion or mood overrides (Fig 1).

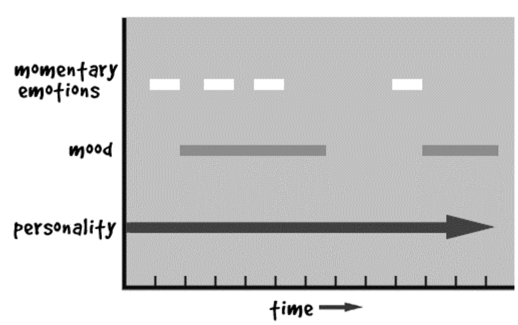

Fig 1.

These behavior levels have different levels of prominence. Reactions or momentary emotions have the highest priority. If no reaction is generated we generate behavior based on the current mood. If our mood level is below a certain threshold then we generate behavior based on our underlying personality (Fig 2).

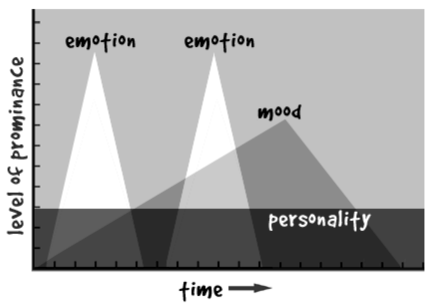

Fig 2.

The characters personality modulates all of the systems in the engine and produces unique behaviors for each character. This allows the engine to be very small, elegant and robust and still provide an almost infinite array of behaviors.

The most powerful behaviors that the engine produces are the character's body and facial gestures. These are very important attributes in the creation of believable interactive characters. These gestures add extra channels of communication to a character and help not only in expressing emotion but also help to push the viewer's brain over a "believability acceptance threshold". These behavioral gestures can be broken down further and described in terms of specific body parts. Facial gestures are comprised of eyebrow movements, eye movements and mouth movements. Body gestures are comprised of spinal movements, shoulder movements, head and neck movements and hip movements.

In the following sections we will describe the model that the engine is based upon, the architecture of the engine, the engines interface (API), the different modules of the engine and we detail an example scenario of the engine at work and how it fits into an application.

A Model of Human Personality.

The conceptual model we use here is based on the extensive work of Professor Hans Eysenck (1985) and Professor Jeffery Gray (1970,1971,1975) who are two of the most prominent researchers in the field of personality psychology and neurology. Their own models represent some of the most robust and solidly grounded work in the field. The model takes fundamental elements from Eysencks behavior and psychobiology based top down model and Gray's neurological and psycho-biologically based bottom up model and combines the two to form a new synthesis representing the emotional brain. Although this model was developed with the intention of being used to simulate rather than authentically reproduce human personality it has, none the less, been described by Professor Gray as a valid and sound model.

What is personality? We define personality as the genetic and environmental differences in the brain structures of individuals. As a species our brains are almost identical which gives rise to our common sets of behaviors. We are, however, all genetically and environmentally unique and as such our brains, although similar, are all different to varying degrees. It is these differences that give us our unique behavioral variations to common behavior patterns, our personality.

The model defines a characters personality as an area in 3 dimensional Cartesian space. The axes of this space are **E**xtroversion, **F**ear and **A**ggression (EFA space). Personality traits are represented by points within this space and are positioned according to the amount with which they are correlated with each axis. For example the personality trait of anxiety is positioned at (**E** -30, **F** +70, **A** -10). This tells us that the trait of anxiety is associated -30% with Extroversion, 70% with Fear and -10% with Aggression (Fig 3).

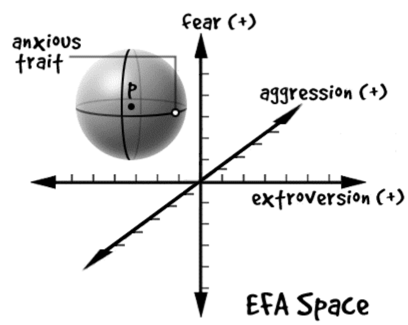

Fig 3.

The position of the center of the personality area in this EFA space (point P) represents by how much each of those axes (central traits) define the overall personality. From this personality point we produce a sphere which will contain the set of personality traits that make up the aggregate personality and are available to the character. Representing the many subtle facets of human personality as being constructed from only three constituent parts seems on the surface like a gross over simplification. However if we only partition our EFA axes into intervals of 10 units we have a potential 20 x 20 x 20, 8000, different personality traits. In general the average person would find it difficult to define more than 50 individual traits. Think of the way in which the simple colors of red green and blue can be combined to produce millions of colors and we have a good analogy of how the three dimensions of personality combine to produce many of the facets of human behavior.

The reason for having three central dimensions of personality becomes clear when we consider that the brain may have three central systems that mediate behavior. We believe these are the Approach System which is associated with **E**xtroversion, the Behavioral Inhibition System which is associated with **F**ear and the Fight / Flight system which is believed to be the biological basis of **A**ggression (Gray 1975). The behavior we display at any time is thought to be controlled by these three systems and the way in which we behave is determined by the genetic prominence of each system. So if our personality were that of an anxious person then our Behavioral Inhibition System would be very prominent in determining our behavior, our personality would have a high Fear component and we would be generally fearful or cautious.

A Model of Human Moods

As with most elements of this system mood is determined by perceived signals of punishment and reward but modulated by the characters personality. This means that the characters personality position in EFA space affects the range of positive and negative moods and also the moods rate of change. The maximum level of positive moods increases with increased **E** , the maximum level of negative moods increases with increased **F** and the speed at which moods build up and decay increases with increased **A** values. If we take a character with a personality that was in the high **E**, high **F**, high **A**, area of our 3 dimensional space their moods would display large positive and large negative values and would build up and decay (change) quickly. So our character would be very "moody" and have large mood swings, as is the case for people with personalities in this area.

Artificial Emotion Engine Architecture.

As we can see from figure 4 the engine is built around nine modules. Six of these modules represent conceptual neural systems, Emotional Reactions, Personality, Punishment and Reward, Mood, Memory and Motivations. Input and Output represent the engines interface, the API. The Self State is the central data repository for the engine and represents the characters general emotional state at any time. The arrows in the diagram represent the feedback and interactions within the system, making it highly dynamic and complex in nature and allowing for some degree of self-organization.

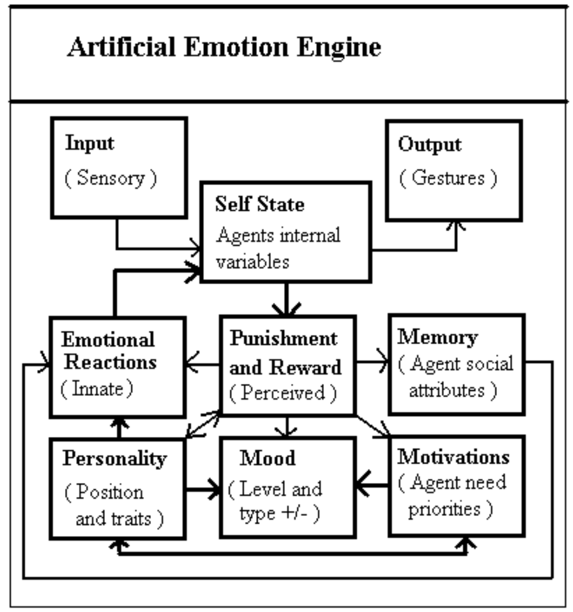

Fig 4.

Punishment and Reward. We, like most animals, seem to function with just two types of input signals, punishment and reward. To this end the majority of this system is fueled by these two inputs wherever possible. When it is not possible (i.e. it is not yet understood) we use higher level conceptualizations. The function of the Punishment and Reward module is to take the incoming signals of raw, sensed punishment and reward (p/r) and translate them into *perceived* signals of p/r. The difference between these two varieties of p/r is that the perceived p/r is dependant on the characters previous history of received p/r and it's personality. Two concepts are used here, habituation and novelty. The more the character receives, in succession, a signal of one type, the lower the effect that signal has. This is habituation. Conversely, the longer the character goes without receiving one type of signal the greater the effect of that signal when it is received. This is novelty. The characters personality determines how susceptible they are to signals of punishment or reward. For example the psychopath is highly susceptible to reward but highly unsusceptible to punishment. This makes them go after 'thrills' without regard for the consequences.

Motivations. Our motivations are arranged into a four layer hierarchy of needs (Maslow 1970). The physiological needs are at the bottom followed by safety needs, affiliation needs and finally esteem needs. Each of these layers is made up of specific sub needs. The physiological need layer is comprised of the hunger, thirst, warmth and energy sub needs. Punishment and reward determines whether each of these needs has been satisfied or frustrated (is closer to or further away from an equilibrium point). We prioritize these needs depending on how far they are from their equilibrium point and this determines what action the character needs to take to rectify the situation. You may have realized that although this engine does not have any cognitive components these motivations supply a very wide range of 'low level' actions for the character to take. These actions are at the level of lower order primate behavior and may suffice for many applications. Personality determines where each layer sits in the hierarchy. Physiological needs are always the top priority but the other layers are interchangeable. For example our psychopath may prioritize esteem above affiliation (friendship and kinship) depending on how much of a psychopath they are.

Personality. The description of the model of personality gives a good conceptual view of this module. Basically it represents a database of personality traits and their position in EFA space along with a rating of the traits social desirability. This 'social desirability scale value' determines whether the trait should be used when the character is in a good (+sdsv) mood or a bad (-sdsv) mood.

Mood. The mood module takes the incoming perceived signal of punishment or reward and uses it to alter the characters current mood. Obviously punishment pushes the characters mood in a negative direction while reward does the opposite. How far and to what level the mood alters in any one time step is a function of the personality (as mentioned previously). The psychopath, with a large **A** component to his personality has moods that change quickly and so is easily pushed into a negative mood by a few signals of punishment.

Memory. In this version of the engine we are implementing a very constrained form of memory to facilitate some non-cognitive social processing. It is therefore a 'social memory'. It is used in conjunction with the affiliation and esteem needs to provide enhanced social behavior, which is the primary application focus of this engine. It will keep track of how much a particular character is liked based on a variety of factors and the social preferences of our family and friends amongst other things. It will also allow us to deepen the abilities of the reactive emotion module using this enhanced data.

Reactive Emotions. These are, generally, innate emotional reactions that have not involved any deep cognitive processing. As such they represent stimulus response reactions of the kind hard wired into our neural circuits. We use the six types of emotional reaction described by Ekman as Joy, Sadness, Fear, Anger, Surprise and Disgust. Again, personality plays a part in modulating these reactions. Our psychopath would show very little fear in his reactions but may show a great deal of anger. The signals of punishment and reward and the levels of the motivational needs trigger the reactions if these signals cross a reaction threshold.

The Self State. This is where we centrally store all data about the emotional state of the agent.

API Input.

The interface to the engine is designed to be as simple and elegant as possible leaving the developer with very little work. The engine requires only signals of punishment or reward for each of the physiological and safety needs (eight in total) in its input. It also takes input representing signals received by the five senses for emotional reactions. Some of these will remain constant if not changed, like the shelter need (we assume we are sheltered unless we are informed otherwise). Some will increase/decrease unless information is received to the contrary (i.e. hunger will increase unless we receive a reward signal for the hunger need). Obviously the ideal situation would be for the character to determine for itself when it was, for example, sheltered or feeding. But right now, in a general-purpose engine, this is not possible and the (small) cognitive burden has to be placed with the developer/designer. A data structure is provided to hold these values and each (developer-determined) time step a call is made to update the engine. This, apart from the characters engine initialization, is all that is required as input to the engine.

API Output.

Output is provided in three formats to give the developer their choice of power or flexibility. At it's most expressively powerful the engine generates the positional information required to produce emotional body and facial gestures. For example, if we had a character whose body joints were being driven by an inverse kinematics system the emotion engine would provide the joint deviations from a plain 'vanilla' movement to make that movement emotional. So if we had a timid character the spine would be curved backward, the shoulders hunched forwards and the head down. Further to this the engine determines the characters movement speed, style and smoothness. A neurotic character would move quickly in short bursts and in a very 'staccato', jerky fashion.

At the second level the engine generates a 'semantic' action plan which is determined by the characters motivational needs. This plan not only determines what to do but *how*. The 'how' is taken from the current mood and the personality trait selected. As an example if the character was out in the cold rain and had been for a while their temperature (warmth need) would be receiving a constant stream of punishment. This would have the effect of making the mood level highly negative and the warmth need a high priority. If the character had a timid personality then the resulting 'semantic' output may be: "*increase warmth very anxiously"*. Here *increase warmth* comes from the motivational need for warmth satisfaction, *very* comes from the importance of the need and the mood level and *anxiously* is a trait that this particular character possess and is appropriate for a negative mood (-sdsv). The *how* of this action plan adds a great deal of depth to the action.

The third level of output is simply the characters raw emotional state data. This gives the developer the flexibility to use the emotional state in ways not handled by the other two modes.

Example Scenario.

Jane is a 38-year-old mother of two playing an office simulation. She is the boss and she has to manage the well being of her office co-workers. Unlike most games the tasks that the team have to perform are secondary to their social interactions. Different team members respond differently to the same situations that the team encounters. She has to make them all work well together and keep them happy (the goal). Jane starts out by taking a personality test to determine her own personality. The result is passed to the Artificial Emotion Engine of her character which will now display behaviors somewhat similar to her own. She then decides on the personalities of her five co-workers. She chooses an extrovert, a neurotic, a psychopath, an introvert and a character with a personality similar to her own. Sounds like a sitcom in the making!

Day one, Jane's (alter ego's) first day at work. Her character is greeted warmly by the extrovert who smiles a wide smile, the introvert sits at his desk and carries on typing, looking glum and hardly acknowledging her presence. The neurotic looks mildly panicked, she does not like change, her shoulders slump forward and she curls up looking submissive. The psychopath sticks out his chest, shoulders back and fixes Jane's character with a steely gaze as he marches, quickly and firmly, over to her to make his presence known. The character with the personality similar to Jane's, John, looks unsure. He is unable to decide if his new boss is a good thing or a bad one for him. Jane, knowing his personality, can empathize with him so she makes the first move to smile and greet John…

What we are trying to show here is that, in addition to improving current genres, a new generation of interactive entertainment experience based on emotion and social interaction is possible using this kind of technology.

Conclusion.

The use of a solidly based, thoroughly researched and credible model of personality and emotion has reaped many benefits with the work we have done. All of the systems that have built on this core model, moods, gestures, behaviors etc. have fitted neatly into the overall scheme and convince us further of the validity of the model. It has given us a solid foundation from which to build more complexity and depth for our characters. Recent research and development has centered on the neurological basis of the three axes of the personality model and their function in a system of emotion. Using the work of Joseph LeDoux we have been able to add non-cognitive emotions on to our model of personality. Autonomous and semi autonomous characters can be given more depth and react emotionally to world and designer scripted events. These characters will help populate the new frontiers of the Internet and virtual environments that we are building.

Future Work.

Simulating human emotion is an endeavor that is very much open ended, at least for the next twenty years or so. We plan to add cognitive processing of some sort to the system and also migrate the systems modules to a neural network based representation. Further to this we anxiously await the discoveries being made by colleagues in the fields of neuroscience.

References.

Ekman, P. & Rosenberg, E. Eds. 1997. *What the face reveals*. Oxford University Press.
Eysenck, H.J. & Eysenck, M.W. 1985. *Personality and individual differences*. Plenum.
Gray, J.A. 1970. The psycho-physiological basis of introversion - extraversion. *Behavior research and therapy, 8*.
Gray, J.A. 1971. *The psychology of fear and stress*. Cambridge University Press.
Gray, J.A. 1975. *Elements of a two process theory of learning*. Academic Press.
LeDoux, J. 1998. *The Emotional Brain*, Touchstone.
Maslow, A.H. 1970. *Personality and motivation*. Harper and Collins.
Wilson, I. 1996. Believable Autonomous Characters. Thesis. Dept. of Artificial Intelligence, Univ. of Westminster.