Emotions in Intelligent Agents

N Parameswaran
School of Computer Science and Engineering
University of New South Wales
Sydney 2052
paramesh@cse.unsw.edu.au

Abstract

In this paper we propose a computational model for implementing emotions in intelligent agents. Emotional behaviour is viewed as a complex mental behaviour directed towards a set of objects in response to changes in the agent's environment. This mental behaviour is responsible for generating different types of mental objects: plans, goals, attitudes, and sub-emotions, which persist in the mental states of the agent for varying lengths of time. We present our preliminary results we obtained by studying emotional agents in a simulated electronic market.

Introduction

Emotions play different roles in controlling the behaviour of individuals in a society. It may have a selfish motivation, or may have an altruistic motivation. For a society to exist in a stable manner, individuals need to display both types of behaviour.

In particular, there is a bi-directional interrelationship between the emotions of an individual and the social norms with emotions playing an instrumental role for the sustenance of social norms and social norms being an essential element of regulation in the individual emotional system. In this paper we have proposed a computational model for emotions, which are displayed by agents during their problem solving behaviour. We present a simple architecture for an agent and the details of an implemented a society of agents using this architecture. We finally discuss our preliminary results obtained in a simulated e-commerce world where the agents display emotion-based behaviours during their shopping activities in the e-commerce world.

Finally, it is argued that the present effort indicates a promising attempt towards incorporating emotions as well in the problem solving behaviour in addition to including the traditional reactive and deliberative behaviours.

Emotions

Emotion plays an important role in many of the human activities both in direct form and in indirect form. Often, a particular human action or utterance of a sentence is understood in a chosen emotional context. Though the field of Artificial Intelligence has been promoting decision making processes based on rational reasoning constrained by consistency criteria until now, humans are known to easily violate rationality and consistency criteria in their day to day lives. One of the occasions when humans do this occurs when they become emotional. Rationality and consistency suffer when humans switch from one emotional state to another. Thus, for example, a well-organized kitchen may become highly disorganized when a housewife starts throwing away the washed vessels in an apparent anger. Though such emotional behaviours often do not appear to contribute to achieving any goals of the agent, they do play an important role in changing one's own mental states. Thus, an upset housewife after throwing away the vessels may feel that her anger is all gone or that she is now satisfied that she has let her husband know how angry she felt, etc. In this paper, we have viewed emotion as an *unplanned* complex mental behaviour.

An emotion generates a behaviour that is similar to the behaviours generated by complex attitudes in human beings. We thus, view emotions as *implicit* attitudes having certain purpose as opposed to explicit attitudes that the agents may "consciously" possess. To keep our treatment simple, we restrict our attention to only those kinds of emotions that produce either favorable outcome or unfavorable outcome with respect to a specific object, such as anger towards an *agent*, ***desperation*** for *food*, happiness due to an *event*, etc. We also assume that the agents we discuss exist with other similar agents in a society.

An emotion with respect to an object, in an agent thus may be defined as a *predisposition to respond in a consistently favorable or unfavorable manner with respect to that object*. Thus, an agent that is under the influence of an emotion may perform different behaviors with respect to an object at different points in time, but these behaviors must exhibit an overall consistency with regard to the implied purpose of the emotion. We discuss emotions and behaviors below.

Emotions and Behavior

An agent under the influence of an emotion will in general display three types of behaviours: physical behaviour, communicative behaviour and mental behaviour. By physical behaviour we mean the behaviour that results when the agent executes a sequence of physical (that is, world) actions. Similarly, communicative and mental behaviours result when communicative actions and mental actions are executed, respectively. Of these three types, the mental behaviour under the emotional influence is the most interesting one, and we focus on this aspect of behaviour in this paper.

Attributes of Emotions

We have already said that an emotion exists with respect to an object.

Directedness

Emotions are polarized; that is, they act in certain biased ways with regard to certain objects.

Consistency When an agent adopts an emotion ***E***, the agent has to exhibit a behavior considered appropriate for that emotion. In a dynamic world, this behavior must include appropriate responses to all situations including unexpected state changes, failures of current activities, and changes in other agents' mental and physical behaviors. Such behaviours typically occur over prolonged periods of time. Thus, it is necessary that the set of sub-behaviours are consistent with the over all motive of the emotion. Thus, for example, an agent ***A*** which is under the influence of **anger** with regard to an object *x* must not involve in a behaviour that may lead other agents to conclude that the agent ***A*** is *not angry* about the object x.

Persistence and Change

An emotion ***E*** once "adopted" by an agent must persist for a *reasonable* period of time, ***T***. Thus, the mental behaviour (which involves generating goals, plans, commitments, and intentions) must be so selected that the physical and communicative behaviour expected of the agent by the other agents in the society persists over a reasonable period of time ***T***. The duration of the period ***T*** itself is a socially defined quantity. Intuitively, the value of ***T*** must be large enough so that the agents interacting with the given agent ***A*** will be able to infer (perhaps through direct communication) or recognize (perhaps through any mental state recognition algorithm) that the agent ***A*** is indeed in the emotional state ***E***. Further, the extent to which the agent is committed towards holding on to a given emotion in dynamic worlds must also be included in the behavioral specifications of the emotion ***E.***

Abstract Emotions

It is possible that when an emotion directed behaviour becomes complex, it will have several subcomponents where each sub-behaviour may correspond to some lower level emotions. This observation indicates that several emotions can be grouped together to define an abstract emotion. Let ***E*** be an emotion with respect an object $\boldsymbol{x}$, denoted as $\boldsymbol{E}(x)$. Let $x_1,...,x_n$ be the sub components of *x*. The emotion ***E*** then can be defined to consist of two parts: $\boldsymbol{E}_1(x_1),...,\boldsymbol{E}_n(x_n)$ where $\boldsymbol{E}_i$'s are (sub)emotions; and emotion ***E'*** is the emotion directed towards the fact that the object x is composed of the sub objects $x_1,...,x_n$. The sub-emotions $\boldsymbol{E}_i$'s produce the necessary sub-behaviors; and the behavior corresponding to ***E'*** takes into account of the fact that the object $\boldsymbol{x}$ is composed of $\boldsymbol{x_1,...,x_n}$. For example, consider an agent that remains ***annoy***ed for the duration $\boldsymbol{T_1}$, and ***angry*** for the duration $\boldsymbol{T_2}$. Then, we may abstract out these two emotional behaviours and represent it as ***desperate*** over the interval $\boldsymbol{T}$ where $\boldsymbol{T} = \boldsymbol{T_1} + \boldsymbol{T_2}$. Thus, emotional behaviours are not necessarily reactive and primitive always.

Multiple Emotions

Sometimes, an agent may possess more than one emotion at the same time towards a given object. The behaviours corresponding to these emotions will involve interleaving and executing the actions from the multiple behaviours. Note that this involves considerable amount of reasoning about the actions that are being scheduled and executed in order avoid conflicts that can potentially arise. For example, it is possible that an agent may be both ***angry*** and ***desperate*** at the same time about a given object.

Representing Emotions

An agent can hold an emotion towards any given object. The object may be a world object such as a house or car, or a mental object such as a goal, a plan, or an agent. For example, an agent may be happy about a car, ***unhappy*** about a goal or plan, and ***angry*** with another agent.

An emotion towards an object $\boldsymbol{x}$ in the world has several attributes (see example below).

Example This example illustrates the emotion ***panic*** of an agent with regard to an object in an e-commerce world where the agent is buying an item in a dynamic market.

Name of the emotion: ***panic;***

Description of object: items at hand to sell;

Basic agent behavior with regard to the objects: Attempt to buy the item as ***quickly*** as possible. Note that ***quickly*** is an attitude. One way to achieve this is to send request to many shops, increasing the price offer each time. Continue this behaviour indefinitely irrespective of what happens in the world, unless the emotional state changes from ***panic*** to some other value.

Consistency: The background computation must make sure that during the ***panic*** driven behaviour, the agent is not assigned any other behaviour that will not be consistent with this emotion. For example, elaborate negotiation during bargaining in order to maximize the profit will be considered inappropriate for ***panic*** emotion.

Persistence of emotion: One way to implement persistence is to enumerate the conditions under which the agent may continue to stay in the given emotional state. In the current example, if the agent did not buy any items over the past several units of time, it may continue to stay in panic state. (However, when most of the objects are sold out, this emotion may be dropped.)
Concurrent emotions If the agent's emotional state did not change within a reasonable period of time, the agent will adopt an additional emotion called ***fear***.

Evaluation Before the agent changes to the emotional state ***panic***, the background computation of the agent must make sure that the change of emotion is justified. For instance, in the example above, the following conditions may be checked to see if they are satisfied:

- Are items to be sold urgently?
- Are there too few agents making offers for these items?

If these conditions are satisfied, the agent's emotional state is changed from its current state to ***panic***. Notice that for simplicity sake, the change to panic did not take the current emotional state into account.

Agent Architecture

The architecture of our agent has five important modules: a Goal Generator, an Emotion Synthesizer, an Attitude Generator, a plan structure module TLS (called the Time Line Structure), and an Executor. The agent continuously monitors the world and updates its world model. The Goal Generator generates goals in response to changes in the world and the messages received from other agents. The planner derives plans for the goals and loads them on the time line structure. The time line structure is a structure where plans that are currently active are loaded and executed, a few actions at a time in the agent cycle. The Emotion Synthesizer synthesizes emotions appropriate for the current situation, and binds them to the goals, and plans. The action executor executes the actions in the time line structure. When the executor encounters an action or a plan with an emotional value bound to it, the *emotions are translated into attitudes and the attitudes are then bound to the nodes* in the time line structure. Once emotions are thus converted into a set of equivalent attitudes, the executor will interpret the attitudes so as to generate the behaviour as required by the corresponding attitudes. Execution of action by the action executor produces changes in the world, which are then sensed by the agents and the other coexisting agents.

Emotion based agents in e-commerce applications

We implemented emotion based migrating agents in a simulated e-commerce application where the market was populated by 20 migrating agents with varying types of behaviours over 20 shops distributed over 20 servers on the network. The migrating agents were implemented using *aglets*. A market directory used in the experiment contains the information on the shops location (URL) available in the market. This is chosen (instead of telling the agents exactly where to shop), so that the traversal behaviours can be investigated. The shopping agents have to self-explore the shops to find out where the tasks can be done. Each shop is implemented so that it can only handle one transaction at a time. The performance of each type of agent ***A*** was investigated in four different type of environments. Three types of behaviours were considered: ***panic***, ***careful***, and ***normal***. An environment consisted of 20 background mobile shopping agents. We considered four types of environments: a) all background agents were panic; b) all background agents were careful; c) all background agents were normal; d) background agents had mixed types of behavior – some were panic, some careful, and some were normal. ***Panic*** is an emotion, and ***careful*** is an *attitude*. ***Normal*** agent is neither emotion driven nor attitude driven.

Three parameters were measured over several experiments and they are: task success rate, time taken to complete the tasks, and the number of migrations (hops) made while achieving the tasks. Task success rate refers to the number of tasks completed relative to the number of tasks assigned. This parameter is used to study the efficiency of the shopping agent. Time refers to the amount of time taken to complete a give task. Time is a sensitive function of the task deadlines. As each type of behaviour differs in the way items are bought, investigating time can give interesting results.

Traversal effort refers to the number of migrations a shopping agent makes before completing the tasks. It is calculated as the ratio between the number of hops (move from one host to another) and the number of tasks assigned. With each migration, the aglet code is transported over the network thus accounting for a portion of the network load. Therefore unnecessary migration should be minimized when this parameter is included in estimating the overall performance. Since this paper focuses on the mobility of the agents, this parameter will reflect how efficient each shopping strategy (traversal) is.

Panic agents: These agents traverse across the market, and buy at the first shop, which satisfy the task constraints. These agents do not bargain. They also do not use the knowledge they have about the market and consequently they do not plan their activities. These agents also ignore the deadline and budget constraints given to them.

Normal agents: These agents traverse and buy only those items that have the closest deadlines at anytime. The immediate availability of the items as they hop over the network is of no concern to these agents. However, they do choose which item to buy next and plan their activities. Buying also involves bargaining, as the agent is not in a panic mode.

Careful agents: These *attitude* driven agents are the most careful of all other types. They travel all shops, collect product information, plan itinerary, and execute the plan. The initiate bargains on selected items for which deadlines are not too close.

Simulation Results

We now present our preliminary results.

Panic agents tend to complete the tasks assigned to them much faster than the other types of agents. Because the tasks are most constrained by deadlines and budgets, desperate agents are fast in buying and have the highest task success rate under our testing conditions. Also, they always make fewer hops, as some tasks would be completed in the same shop. While emotional agent appears to have outperformed the other two types of agents, they have made a trade-off in risking the best buys. When the same product is sold at more than one shop in the market, the agent's behaviour did not guarantee the cheapest buy.

Careful agents show good performance in all testing parameters except in time. It achieves its good performance by making more hops, collecting more information, planning a buying strategy, and finally buying the items. All this involves more time, thus affecting the speed of the agent. The task completion rate is high, comparable to the emotional agents. Though the agents' activities were slow at the shops and they traveled more across the network, the items were bought at their cheapest prices.

Normal agents show relatively poor performance in all our tests. They generally take the longest time to complete a given task, which is reflected in the way they go about doing their business. The deadline for each task was chosen so that it accommodates the average time required to complete each task for the **panic** and **careful** agents. This resulted in the deadlines of tasks that are relatively closer to each other. **Normal** agents consequently encountered problems as many tasks are due at around the same time, and they consequently failed to complete all the tasks.

Background agents

In all the above experiments, we investigated the behaviour of a given agent ***A*** under a chosen type of background agents. The performance of the agent ***A*** remained fairly unaffected when the type of the background agents was the same as that of ***A***. Thus, if ***A*** was **careful**, its overall performance was not affected by the shopping behaviour of the background agents when background agents were all **careful**. However, when the background agents were of a different type, the performance of all types of agents deteriorated except for the panic driven agents. For the ***panic*** driven agent, the performance was fairly stable most of the time in all types of background agents. The worst affected agents were of the type **careful**. The **normal** type agents showed a performance that appeared to be a compromise between other two types.

Related work

(Breazeal 1998) reports an implementation of a robot, which showed emotional expression when a human interacts with it using a GUI. This robot was mostly reactive in nature and displayed emotional expressions anger, interest, happiness, disgust, surprise, fear, sadness, and excitement. The emotional response is programmed to be non-conflicting in nature. However, emotions did not produce any long-term behaviour while in our model they do. (Ushida 1998) proposes an architecture, which is based on the cognitive appraisal theory from psychology. In this, emotions are synthesized using a set of rules after deliberating over the events that have presently taken place. The agents respond emotionally by generating simple goals, which are then achieved using a simple planner. Thus, the emotions affect the agent behaviours by directly affecting the agent's actions unlike in our approach where the emotions affect the actions as well as the *attitudes* thus emotions leaving a permanent or semi permanent mark on the agent behaviour. In (Velásquez 1998), emotions are viewed as biasing mechanisms in the decision making process inside the agent. Many other works on emotions is focused on entertainment. Other important works include (Reilly 1996) and (Elliot 1992). They mostly deal with contribution of appraisal to emotions and thus emphasizing cognitively generated emotions as we do.

Conclusion

We have considered in this paper only those types of emotions, which are *cognitively generated*, *directed* and exist *with respect to an object*. However, we notice that humans display more complex emotions, which are not necessarily directed and are not object centered. For example, humans can feel happy for unknown reasons without involving any object. A comprehensive model for emotion will thus need to take into such parameters as environment, situations, past and future mental states of the agent under consideration and the other coexisting agents in the society.

References

Breazeal, C. 1998. A Motivational System for Regulating Human-Robot Interaction, Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)

Ushida,H. 1998. Yuji Hirayama and Hiroshi Nakajima. Emotion Model for Life-like Agent and Its Evaluation, Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)

Velásquez, D.1998. When Robots Weep: Emotional Memories and Decision-Making Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)

Staller, A. and Paolo Petta. 2001. Introducing Emotions into the Computational Study of Social Norms:A First Evaluation, Journal of Artificial Societies and Social Simulation vol. 4, no. 1, http://www.soc.surrey.ac.uk/JASSS/4/1/2.html .

Elliot, C.D. 1992. The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System. Ph.D. Thesis, Institute for the Learning Sciences, Northwestern University.

Reilly S. 1996. Believable Social and Emotional Agents. Technical Report, CMU- CS-96-138, School of Computer Science, Carnegie Mellon University.