Emotion as a Significant Change in Neural Activity

Karla Parussel

Department of Computing Science and Mathematics
University of Stirling
Scotland

It is hypothesised here that there exist two classes of emotions; driving and satisfying emotions. Driving emotions significantly increase the internal activity of the brain and result in the agent seeking to minimise its emotional state by performing actions that it would not otherwise do. Satisfying emotions decrease internal activity and encourage the agent to continue its current behaviour to maintain its emotional state. It is theorised that neuromodulators act as simple yet high impact signals to either agitate or calm specific neural networks. This results in what we can define as either driving or satisfying emotions. The plausibility of this hypothesis is tested in this paper using feed-forward networks of leaky integrate-and-fire neurons.

Introduction

Driving and satisfying emotions

Emotions in a natural agent are either pleasant or unpleasant, but never neutral, (Nesse, 1990). We can understand emotions as being either positive or negative, (Zhang & Lee, 2009). This may be useful when describing emotions from a personal perspective, but this description carries connotations and an implicit judgement on their utility. We need to differentiate between the experience of an emotion and its effect. For example, it may be argued that anger is a negative emotion because it is unpleasant to experience. An equally valid argument is that anger has positive motivational benefits. Nesse gives an example of the rationality of anger. In a long term, committed social partnership where one party is tempted to defect, the threat of an irrational and spiteful retaliation because of the betrayed partner's anger decreases the likelihood of a defection continuing or even taking place at all.

We could also think of emotions as being either attractive or repulsive. This may be useful when describing emotions within the context of a dynamical system but it is less applicable when describing animal or human behaviour. For example sadness and fear are repulsive emotions that we normally seek to minimise but people deliberately invoke sadness by watching soap operas and other melodramas. They deliberately invoke fear when reading or watching thrillers or by participating in fun-fair rides or extreme sports. In the latter case, people engage in these activities because they are also exciting and fun. Emotions can be simultaneously attractive and repulsive.

Rolls describes emotions in terms of rewards and punishments. An animal will work for a reward, but will work to escape or avoid a punishment (Rolls, 2005, page 118). Emotions are proposed as being states elicited by rewards and punishers and changes in reward and punishment (Rolls, 1999, page 60). Contentment can be considered a rewarding emotional state for example, but how is this different from the emotion of joy? Rolls uses the concept of positive and negative reinforcers and punishers as determined by whether the reinforcer or punisher increases the probability of a response by the agent.

These terms may be useful when describing observations of animal behaviour but they are less descriptive when referring to emotional experiences or an appreciation of why rewards and punishers have the effect that they do. Animals *experience* reward and punishment. From the perspective of the animal, reward and punishment is more than mere habituation and conditioning.

It can be seen that the utility and limitations of the descriptions that we use partially depend upon the context in which they are employed. Can we decide on terms that are unambiguous regardless of whether we are referring to the experience of emotions or observations of animal behaviour? It is proposed here that emotions can be thought of as being either *driving* or *satisfying*. These terms describe both the experience and behavioural effect of being in an emotional state and are also neutral as to its utility. This is more than a mere linguistic exercise, there is a theoretical basis behind these terms.

(Rolls, 2005, page 128), discusses how *taxes* orient an organism towards or away from stimuli in its environment. Phototaxis bends a plant towards a light source for example. An organism may move towards sources of nutrients or away from materials with physical properties detrimental to its health.

Animals need to maintain homoeostasis. Various bodily processes need to be kept relatively constant. Critical resources must be kept replete regardless of the environment that the agent may inhabit. Examples of these resources include levels of food, water and oxygen. Internal physiological variables must also be kept within a certain range. For example, natural agents will seek warmth when it is too cold and try to cool down when it is too hot.

Panksepp describes how sensations generate pleasure or displeasure depending on the homoeostatic equilibrium of the body (Panksepp, 1998, page 164). For example, food

tastes better when we are hungry. Panksepp also discusses the idea of emotional attractors in the brain as reflected by repetitive patterns of electrical activities that are triggered by specific environmental stimuli (pp94).

When discussing homoeostasis we can think of driving emotions occurring when the organism needs to reassert the internal equilibrium of its physiological processes. Satisfying emotions would ensue when equilibrium is reasserted or maintained. Although successful adaptation to an environment is more than a matter of maintaining homoeostasis, the concepts apply equally well to neutral actions and behaviours that have no intrinsic value to the maintenance of the body. For example, animals may be driven to seek out others to breed and socially bond with, and can be satisfied and settle down when they do.

So far we have only judged emotions as being driving or satisfying. If these terms are to be non-ambiguous then we need to be able to determine their classification through quantifiable measurements rather than via interpretation.

The hypothesis

The brain can be understood as a self-organising system, (Kelso, 1995); (Malsburg, 2003). The way that a normal brain functions internally is not directly determined by an external controller. Instead the brain reacts to signals from the agent's senses.

If the brain self-organises then there must be attractors or relatively stable states that it can settle into. When other physical systems self-organise they normally do so as their own internal energy dissipates or is minimised. Patterns can emerge instantaneously in a chaotic system but also disappear again just as quickly. Patterns persist in a self-organising system because there is insufficient energy or activity to break them apart. We can see this happen with crystallisation when a liquid becomes supersaturated as it cools. If the liquid is re-heated then the activity of the molecules is increased and the patterns start to break up.

It was discussed above how Panksepp refers to emotional attractors in the brain as seen by repetitive patterns of electrical activities. The assumption made in this paper is that the brain self-organises by settling into stable states, or attractors, as characterised by a reduction in internal activity. This assumption is consistent with how models of artificial neural networks developed by the author are understood to self-organise.

Feed-forward networks of leaky integrate-and-fire neurons can be made to self-organise by minimising the strength of their input activity. This consequently also reduces the internal activity of the network, (Parussel, 2006). The networks act as minimal disturbance systems. Incoming activation is directed to neurons in the output layer and the action that corresponds to the winning neuron is performed accordingly. If the action has desirable consequences then the appropriate input signals fed to the network are temporarily reduced in strength. Actions that reduce the strength of the input signal have a greater chance of being performed again in the future. If an action does not subsequently decrease the strength of

the input signal then other actions have an equal chance of being performed.

It was shown that the networks can be biased towards exploration using inhibitory neuromodulators in the middle layer. This agitates the network out of any relatively stable state to increase the chance of it exploring other actions, (Parussel, 2006). The networks can also be biased towards exploitation using excitatory neuromodulators at the output layer, (Parussel & Cañamero, 2007).

Emotions can be modulated by altering the levels of neuromodulators in a brain (Kelley, 2005). (Fellous, 2004) argues that emotions provide a multi-level communication of simplified but high impact information. He also argues that emotion can be seen as continuous patterns of neuromodulation of certain brain structures (Fellous, 1999).

The hypothesis made here is that neuromodulators are used to either aid or hamper the brain in minimising its internal activity. It is theorised that neuromodulators are used for this purpose to act as simple yet high impact signals to either agitate or calm specific neural networks. This results in what we can define as either driving or satisfying emotions. Driving emotions significantly increase the internal activity of the brain and result in the agent seeking to minimise its emotional state by performing actions that it would not otherwise do. Satisfying emotions significantly decrease the internal activity of the brain and increase the probability of the agent in continuing its current behaviour to maintain its emotional state.

(Noble, 1997) argues that Artificial Life simulations cannot prove theories concerning the real world. The role of such models is to establish the plausibility of a theory. The theory can then be referred back to the relevant empirical science in order to be proven in the natural world. The plausibility of the above hypothesis given the stated assumptions and observations is tested in this paper.

Method

The system

A self-organising biologically inspired neural network has been developed so as to explore the functionality provided by neuromodulation. The intention is to increase our understanding of emotions by researching the functionality of the mechanisms underlying them.

The artificial life animat concept has been abstracted to provide the simplest possible context for testing the effect of neuromodulation when applied to an artificial neural network. A stimulus-response agent has been created that can neither sense an environment nor be affected by one. The only thing that it interacts with is a body with two resources, labelled "Energy" and "Water" (see figure 1).

Each change in resource level is passed to the agent controller as an input signal. Before being input, they are scaled to the largest increase and decrease that has occurred to each resource so as to be within the range [0,1]. They are then inverted so that desirable changes, such as increases to a resource level, result in a reduced signal to the agent controller.
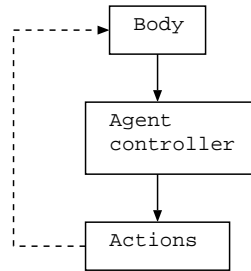
Figure 1. The agent controller receives input signals derived from the state of the body. It then attempts to choose one action to be performed. The action directly alters the body state of the agent. This leads to different input signals being passed to the agent controller in the next turn.

The agent can execute a set of actions that either increase or decrease by a given amount the energy or water level in the body, plus two neutral actions. Neutral actions are useful because if they are used differently to each other then it throws doubt on how well the agent is adapting. The 'Inactive' action is used by default when an agent fails to choose for itself. This can happen if no activation reaches the output neurons of its neural network. It results in each resource of the agent being reduced by the maximum cost. The effect of this is more costly to the agent than if it deliberately chose the most costly action available to it as that would only result in a reduction of one resource.

The neural network

The agent adapts using a feed forward neural network of spiking leaky integrate-and-fire neurons based on the model described in (Koch, 1999, page 339); (Wehmeier, Dong, Koch, & Essen, 1989).

The network learns which outputs should be most frequently and strongly fired to minimise the subsequent level of input signal in the next turn. Each neural network is made up of three distinct layers; input, middle and output.

For each resource, the input layer has two neurons that output to the middle layer. One neuron signals the need for the resource and the other neuron signals the satisfaction of that need. There are situations in which an effective behaviour for an agent may be to decrease a need but not satisfy it. Alternatively there may be situations in which an agent needs to store more resources than it is used to doing. In these experiments the agent is tasked only with maximising its resources.

There is one output neuron per action. The network is iterated over a fixed number of times within a single turn. The action corresponding to the output neuron with the greatest average activation is then performed. The action performed by the agent directly and immediately alters the level of a resource. This consequently determines the strength of the corresponding input signal fed to the network in the next turn. This is fed via the input neurons corresponding to the resource affected by the action. This allows the network to act

as a minimal disturbance system (Wörgötter & Porr, 2004) as it settles upon actions that reduce its total input activation.

The Neuron

Spiking neurons were used in the neural network, each one acting as a capacitor to integrate and contain the charge delivered by synaptic input. This charge slowly leaks away over time. The neurons have a fixed voltage threshold and base leakage which are genetically determined.

The neurons also have an adaptive leakage to account for how frequently they have recently spiked. If a neuron spikes then its leakage is increased by a genetically determined amount. If the neuron does not spike then the leakage is decreased by that same amount [1]. Resistance is constrained within the range [0, 1]. This model was inspired by the adapting integrate-and-fire models described in (Koch, 1999) section 14.2.3 pp339.

The spiking threshold is the same for all neurons in the network and is constant. The neurons are stochastic so that once the spiking threshold has been reached, there is a random chance that a spike will be transmitted along the output weights. Either way the cell loses its activation[2].

The neurons send out a stereotypical spike. This is implemented as having a binary output. The weights connecting the neurons are constrained within the range [0, 1].

Local learning rule.

The learning rule employed uses spike timing-dependent plasticity (Bi & Wang, 2002). It is implemented as a two-coincidence-detector model (Karmarkar & Buonomano, 2002); (Karmarkar, Najariana, & Buonomano, 2002) based on (Song, Miller, & Abbott, 2000) and later evolved for use in robots by (Di Paolo, 2003). Each neuron has its own post-synaptic recording function that is incremented when the neuron spikes and which decays over time in-between spikes. This is compared to the pre-synaptic recording function of the neuron that has transmitted the activation. Each layer of neurons has its own increment and decay rates determined prior to testing via automated parameter optimisation.

Synaptic connectivity between layers.

The multitude of connectivity between two layers is specified using a continuous value whereby the fractional part determines the chance of a connection between two neurons being made. So for example, a multitude of 1.5 would mean that every neuron in a source layer was connected to every neuron in the target layer at least once, but with a 50% chance of being made a second time. Parameter optimisation most often selected multitudes of less than 1.

---

[1] It was not known whether separate increment and decrement parameters were required. To keep the number of attributes to a minimum it was decided that two parameters would be used only if it was found to be required. The network evolved well with only one parameter.

[2] This model of stochastic firing is a simplified one. In real neurons there is both both a possibility that a spike is passed down the axon to the target cells and a possibility of vesicle release once a spike has reached a synapse.

All connections between layers are excitatory and modifiable. Non-modifiable connections were avoided to help minimise the risk that evolution would hard-code the network topology to increase the average fitness during parameter optimisation.

Modulation.

A modulator is a global signal that can influence the behaviour of a neuron if that neuron has receptors for it. The signal decays over time, as specified by the re-uptake rate, and can be increased by firing neurons that have secretors for it.

Neurons that are to be modulated are given a random number of receptors. These can be modulated by neurons in other layers that have secretors for those modulators. The receptors modulate either the neuron's sensitivity to input or probability of firing. The effect of this modulation is determined by the level of the associated modulator and whether the receptor is inhibitory or excitatory.

Neurons can also have secretors. These increase the level of an associated modulator. The modulator re-uptake rate, the modulation rate of the receptors and the increment rate of the secretors is determined by artificial evolution along with many other parameters of the neural network before the model is tested.

Parameter Optimisation.

The parameters of the networks are optimised using artificial evolution so as to make a fair comparison. Once these constrained evolutionary runs are finished the parameters are hard-coded and tested as a population of 450 agents in order to determine the average performance of the neural network. An average fitness is required because the mapping from genotype to phenotype is stochastic. This is due to the randomisation of weights and the connectivity between neurons. The fitness function used during parameter optimisation was $Energy + Water + Age - absolute(Energy - Water)$.

The absolute difference between the energy and water resource is subtracted from the fitness as both resources are essential for the agent to stay alive. Age is important for the fitness function during parameter optimisation when agents are more likely to die before the end of their evaluation.

Minimal disturbance networks

The network learns which outputs should be most frequently and strongly fired to minimise the subsequent level of input signal in the next turn.

It is easier to understand how the neural network functions if it is seen as a dynamical system(D.Beer, 1995). Understanding an agent as a self-organising dynamical system removes the question of when to switch behaviours as the transition happens continuously over time. It also means that it is more appropriate to think of attractive and aversive external stimuli than positive or negative reinforcement. The system can therefore be self-organising and more autonomous. Self-organisation removes the question of when to teach the network and when to recall information encoded in it.

(Wörgötter & Porr, 2004) provide an overview of the field of temporal sequence learning. They discuss how the learning paradigm of disturbance minimisation, as opposed to reward maximisation, removes the problem of credit structuring and assignment. The two paradigms are not equivalent. Whereas maximal return is associated with a few points on a decision surface, minimal disturbance uses all of the points. In a minimal disturbance system, every input into the system drives the learning process. If there is no signal then the system is seen as being in a stable state. Rewards and maximal return are not sought, as is the case with credit assignment learning. Instead, any disturbance-free state is satisfactory.

Minimising free-energy

The dynamics of a self-organising system can be understood using the concept of an energy landscape, (Heylighen, 2000); (Kauffman, 1993, page 176).

Using an analogy of a ball rolling along a peak, ridge or plateau, then given sufficient energy it will roll down a slope and minimise its own potential energy. The ball will not be able to later return unless its kinetic energy is first increased. This process will continue until the ball comes to a stop at the bottom of the landscape, or within a local depression that requires more kinetic energy than the ball currently has for it to escape. Valleys correspond to attractors in a dynamical system, the speed that the system moves into them being determined by the steepness of the slope.

This is not a new concept in neural network theory, an energy function was first used with Hopfield networks (Hertz, Krogh, & Palmer, 1991, page 21). This allows an 'energy landscape' to be imagined whereby patterns memorised, being attractors in the system, can be seen as local minima in the landscape. As with the analogy of the ball, assuming the influence of gravity, a particle placed anywhere on this imaginary surface will roll down to the nearest basin.

Adaptive performance of the networks

The synaptic weights between the input and the middle layer of the network can be thought of as providing 'activity diffraction' to allow the input signals to filter through the system at different speeds. The synaptic weights between the middle layer and the output layer can be thought of as providing 'activity integration', integrating those signals back into combinations that allow particular output neurons to fire more frequently than others.

The network is feed-forward rather than recursive. The output neurons do not connect back to the input layers but they can affect them indirectly. The action that corresponds to the output neuron that has the highest average activation over all the iterations within a turn is performed by the agent. If this increases or decreases a resource in the body then this change is reflected in the subsequent input signals fed to the neural network.

Because activity filters through the network at different speeds, some output neurons will fire earlier than others. If an action is rewarding and subsequently results in a reduction of input signal to the network, synaptic activity will be

reduced for the other neurons and therefore will be less likely to fire. If an action is not rewarding, the input signal is not reduced, other neurons will eventually fire and other actions will be tried instead.

Biasing a network for either exploration or exploitation

Neuromodulators can be used to bias a neural network to function in a certain way depending on how they are used. The networks used here can be biased towards either exploration or exploitation by using inhibitory or excitatory receptors respectively, (Parussel, 2006); (Parussel & Cañamero, 2007).

The network can be biased towards exploration if the hunger and thirst input units secrete corresponding modulators, for which the middle layer units have a random set of inhibitory receptors (see figure 2a). Alternatively, the network can be biased towards exploitative behaviour if the middle layer secretes a single modulator for which the output layer has excitatory receptors (see figure 2b). A variant of the exploitation network that can be externally influenced using modulators can be seen in figure 2c. This network was optimised for use without modulators but was tested with inhibitory receptors applied to its input layers.
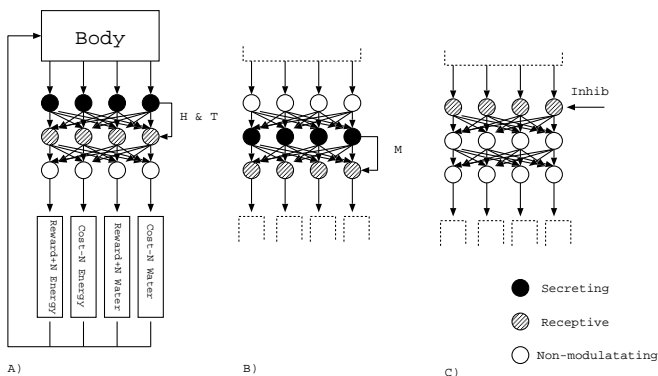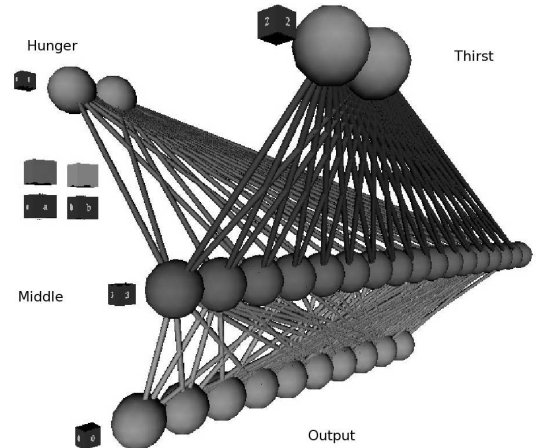

Figure 3. The structure of a typical modulating network. Clockwise from top left; (top left) hunger signal layer, (top right) thirst signal layer, (middle) middle layer, (bottom) output layer. One neuron in each input layer signals the decrease in the corresponding resource, the other signals the resource being repleted. The output layer consists of one neuron per action available to the agent.

Figure 2. The agent has a body that contains water and energy levels. There are four input neurons; two per resource, one to signal satisfaction of the resource and the other to signal the need for it. Only four neurons are shown in the above diagram for the middle and output layers. In practise the exploratory network had fourteen neurons in the middle layer and the exploitation network had nineteen. Each network has ten neurons in the output layer, one for each action. A) Two-modulator agent biased for exploration: Hunger (and thirst) neurons increase the strength of the hunger (or thirst) modulator when they fire. Neurons in the middle layer have a random number of inhibitory receptors for these modulators. B) Single-modulator agent biased for exploitation: Neurons in the middle layer increase the strength of a single modulator when they fire. Neurons in the output layer have a random number of excitatory receptors for this modulator. C) Network with inhibitory receptors at the input layer that can be biased towards otherwise neutral actions by a system external to the network.
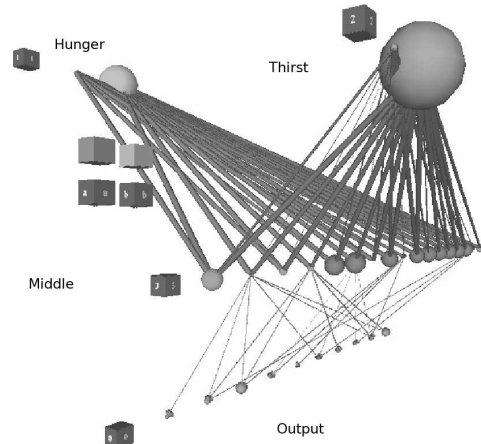
Exploitation.

Figure 4. An exploration network in action. Neuron size shows level of activation, width of weights show synaptic strength. For example in the above image, the middle to output layer connections have low strength and the output neurons have low activation. A neuron in the thirst layer (top right) has a high activation and will probably soon fire. Each upper cube corresponds to a modulator, the height of which is animated to show the changing global strength of that modulator. The non-animated numbered cube underneath is used to identify the modulator; (left box) hunger modulator, (right box) thirst modulator.

With a non-modulating or exploratory network, the more rewarding an action, the stronger the activation of the corresponding output neuron. In contrast, a network biased for exploitation provides reduced activations for all of its output neurons by default and uses modulation to excite the output neurons which are rewarding.

The excitatory receptors at the output layer increase the activity of the network. If the network performs an action

that reduces the overall input level, then the winning neurons in the output layer will benefit more from excitatory modulation than the less activated neurons that are losing.

If the network performs a sub-optimal action then the input signal is reduced less than if an optimal action was chosen. This means that other middle layer neurons are more likely to fire. Each middle layer neuron will also increase the level of the modulator and each is more likely to fire different output neurons. Consequently, the modulator cannot be used to excite any particular output neuron more than all the others when a sub-optimal action is performed.

Exploration.

It was discussed in section how activity filters through the network at different speeds with some output neurons firing earlier than others. If an action is rewarding and subsequently reduces the input signal to the network, activation of the other neurons will be reduced and they will be less likely to fire. If an action is not rewarding, the input signal is not reduced, other neurons will eventually fire and other actions will be tried instead.

The hunger and thirst modulators of the exploration agent optimised for use with discrete actions inhibit the neurons in the middle layer. The strongest firing neurons have more activation to lose when being inhibited. These are also the neurons more likely to be firing the output neurons that lead to actions that reduce total input activity into the network. So by inhibiting the neurons in the middle layer the 'diffraction' of activation throughout the network is reduced and other actions have a greater chance of being performed. This increases exploratory behaviour.

Relevance of exploration and exploitation to emotions.

Emotions can help the reasoning process (Damasio, 1994). Evans puts this idea in a game-theoretical framework in his search hypothesis (Evans, 2002). A rational agent confronted with an open-ended and partially unknown environment, emotions constrain the range of outcomes to be considered and subjectively applies a utility to each. The search hypothesis can be seen as an example of an agent moving from exploration of possible outcomes to an exploitation of the action providing the current expected highest expected utility.

However, the best course of action does not need to be learnt through experience. (Nesse, 1990) defines emotions as specialised states of operation that give an evolutionary advantage to an agent in particular situations. (LeDoux, 1998) describes a distinguishing characteristic of cognitive processing as flexibility of response to the environment. Emotions provide a counter-balance to this by narrowing the response of an agent in ways that have a greater evolutionary fitness.

As an example, predator avoidance driven by fear is an ideal behaviour to be selected for and optimised by evolution. It is a behaviour that needs to be maintained until the prey reaches assured safety regardless of whether it is able to continually sense the predator or not (Avila-García & Cañamero, 2005). Nor will the prey benefit from being distracted by less important sensory input while it is still in danger. Successful

fleeing behaviour might not require exploration of different actions when instead exploitation of known strategies for a successful escape should be given priority. On the contrary, positive emotional states are thought to promote openness to the world and exploration of new courses of actions (Blanchard & Cañamero, 2006).

Results

Changing the level of overall activity

In dealing with the networks described here, it has been observed that it is not as important whether receptors have an excitatory or inhibitory effect on a neuron so much as whether release of modulator increases or decreases activity for the entire network.

The exploitation network has the middle layer excite the output layer when modulating the sensitivity to input of the neurons. Yet when modulation is decreased network activity is increased. Conversely, the exploratory network has the input layer inhibit the middle layer via modulation, yet this decreases overall activity (see figure 5). This graph shows the average activation for all the neurons for both networks and how it either increases or decreases as the modulation rate is increased. The change in the activation of the exploitation network can be seen more clearly in figure 6.

Biasing the network towards exploration requires that activity is increased so that other outputs have a greater chance of winning. Biasing it towards exploitation requires that the network is led to a more stable state by reducing its overall activity. The network controller can also be influenced to choose specific actions that would probably otherwise not be selected by an external system decreasing its overall activity.

Using the analogy of the energy landscape again, it is useful to think of a ball rolling around a local minima. Shaking the landscape via an earthquake may bounce the ball out of it and elsewhere. The stronger the earthquake the more chance there is of this happening. Stopping the earthquake all-together allows the ball to come to a rest. The network is the ball, constantly trying to come to a rest. The modulators are like a gain control for the earthquake.

Decreasing overall activity.

Figure 6 shows the effect of increasing the modulation rate for the exploitation agent. The effect is not as dramatic as for the exploration agent as only the output layer is modulated. The output layer does not connect to any other layer and so the effect of the modulation is localised. Whereas with the exploration agent, the middle layer is modulated and this also has an effect on the output layer.

It may seem strange that the activity of the entire network can be reduced by exciting the output neurons, whether by increasing their probability of firing or by increasing their sensitivity to input. The neurons in the output layer do not connect to any other neurons and so the increase in activity does not affect any other neurons. But because excitatory modulation makes the output neurons more likely to fire, then the neurons are more likely to enter into a refractory state for
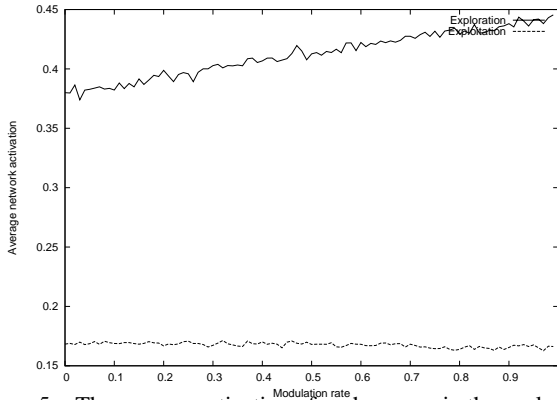
Figure 5. The average activation of each neuron in the exploration and exploitation networks as the modulation is increased. The figures are derived from a population of 450 agents running a network consisting of the same number of neurons for 1,000 turns, with each turn consisting of 10 iterations of the network. The exploitation network referred to here has receptors that excite the neuron's sensitivity to input.
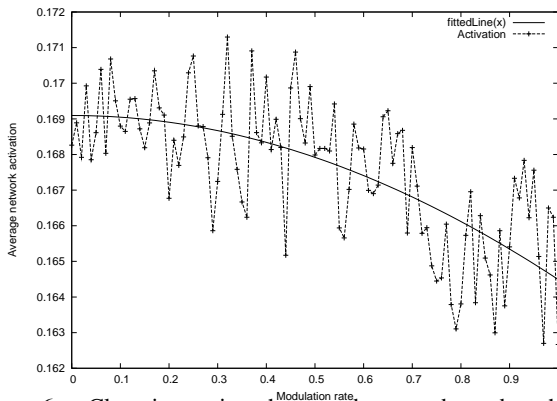

Figure 7. Non-Modulating agent run with inhibitory receptors at the input layer. The *x*-axis the number of receptors that were given to each neuron in the input layer. The *y*-axis denotes the frequency of each action chosen during that run. A population of 450 agents were tested for each run for 1,000 turns.

Figure 6. Close inspection shows a downward trend to the average activation of the exploitation network as modulation is increased. Least-squares fitting using a Lorentzian function has been performed using Gnuplot's fit command.

a period of time. During this period, any incoming activation is immediately leaked away.

The exploitation network uses excitatory receptors at the output layer in order to reduce network activity and to bias it towards performing particular actions. A non-modulating network was also adapted to reduce network activity, but with inhibitory receptors at the input layer instead. The modulator for these receptors was released externally to the network when specific actions were performed. This was used to bias the network to perform two otherwise neutral actions (see figure 7).

Increasing overall activity.

Biasing a network towards exploration can also be used to stop activity dieing all-together. With a network evolved to work without modulators, a decrease in input activity eventu-
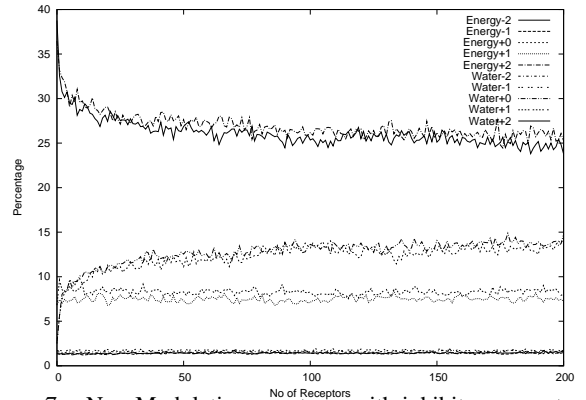
ally leads to a decrease in output activity. If an action results in the lowest possible strength of input signal and spiking-activity in the network has already declined to the minimum threshold required for hebbian learning to occur, then the network settles into a stable state. This can occur in the absence of any changes external to the system, such as the effect of an action changing or noise being added to the input signals.

If no activity reaches the neurons in the output neuron then the agent cannot choose an action for itself. In this situation the default 'Inactive' action is chosen for the neural network controller by the encompassing system. This function is more costly than if the agent chose an action itself.

When testing a population of non-modulating agents for longer than 1,000 turns, spiking-activity in the network would cease over time, (Parussel & Smith, 2005). This led to the weights freezing because the STDP learning rule only updated the weights when spikes occurred. The activation of the output neurons would slowly decay over time with the winning action remaining the same in the absence of any change in the effect of that action (see figure 8). The limited use of artificial evolution for parameter optimisation had settled upon a brittle strategy which depended on how long each agent was evaluated for.

A population of modulating exploration agents were then tested for the same extended period of time. They were shown to continue transitioning between the same two winning output neurons that caused a maximum increase in energy and water, with other neurons very occasionally being chosen (see figure 9).

When comparing the exploration agent and non-modulating agent in environments that they were not evolved for, in this case evaluated for a variable or extended length of time, then it is shown that modulation makes the agent more robust. This robustness carries with it a performance cost as the exploration agent can not execute the actions that only reduce the overall input signal the most.
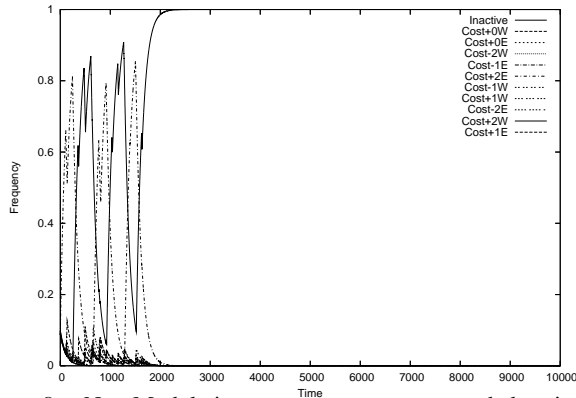
Figure 8. Non-Modulating agent run over an extended period of time (10,000 turns). The *x*-axis denotes the number of turns that the agent was evaluated for, at each turn the agent could choose a different action. The *y*-axis denotes the frequency of that action. Choosing an action increases its frequency resulting in the decrease of frequency for the other actions. The agent quickly ends up alternating between the two actions that provide the maximum increase in energy and water. After about 2,000 turn the agent settles on only performing the action that provides the maximum increase in the water resource even though energy is just as important to its survival.
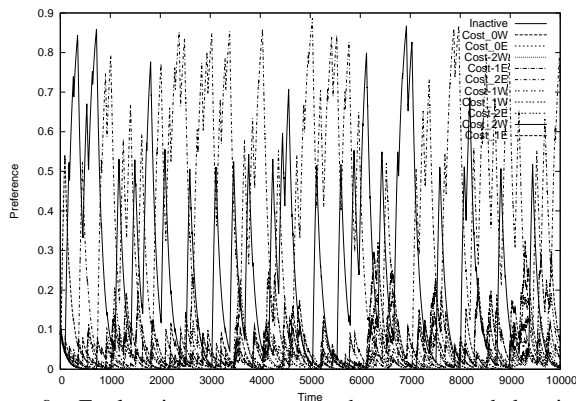

Figure 9. Exploration agent run over the same extended period of time and using the same axes as in figure 8.

Modulation for the exploration agent, stops the system from settling into a stable state where activity declines to a point whereby the network stops alternating between actions. This agitation is achieved by increasing the global network activity of the system. This stops it reducing the strength of its inputs and settling into a stable state, even though the environment may allow for it or make this the optimal behaviour.

Discussion

It has been hypothesised in this paper that there exist two classes of emotions; driving and satisfying emotions. The model of emotion used here is presented as a mechanistic process within the context of a dynamical system The brain,

performing action selection by self-organising and minimising its internal activity, is either hampered or aided in this by neuromodulators that agitate or calm it. This has the effect of either influencing an agent to perform actions it would not otherwise do, or continuing its current behaviour. Because this is a significant influence upon the behaviour of the agent, we observe these effects as driving and satisfying emotions.

Applying the hypothesis to natural agents

The plausibility of this hypothesis has been tested in this paper using feed forward networks of spiking integrate-and-fire neurons. The model is a simplistic one that adapts to immediate stimuli. There is no capacity for long term memory storage, adaptation to sequences of inputs or ability to associate stimuli.

But for a stimulus-response agent, the results are consistent with the hypothesis. This in itself is insufficient for us to draw conclusions about the nature of emotions in animals and humans. For that, we will need real-life observations. There is evidence though to suggest that the hypothesis may apply to natural brains. Panksepp provides some evidence for this when discussing the localisation of activity in the brain related to emotional states, (Panksepp, 1998, page 95-96). Activity is decreased in the brain when happy feelings are experienced, conversely activity is increased when sad feelings are experienced. People who have a predilection towards panic attacks exhibit over-activity in their right parahippocampal regions.

If the hypothesis is correct then it will most likely not be as simple as positive emotional states being signified by low neural activity and negative emotional states signified by high activity. Panksepp provides the example of people suffering depression exhibiting less arousal of their left frontal areas than is considered normal.

(Gotlib & Hamilton, 2008) provide a review of the literature concerning the neural activity of people suffering from depression. They also discuss the increased activation of the amygdala in those prone to depression. Major depressive disorder is a complex phenomenon which is essentially a psychiatric disorder of the regulation of emotion. A lack of motivation and the inability to regulate the processing of negative experiences are strong characteristics of depression. Their review discusses the idea of depression being caused by an imbalance in the activity between the limbic system and the dorsal cortical structures. Also noted is the apparent normalisation of activity between the amygdala, anterior cingulate cortex and the dorsolateral prefrontal cortex following successful treatment. Minsky raises the issue of competing interests within an artificial agent giving rise to what can be observed as emotional reactions (Minsky, 1988). Maybe a related point is that emotional disorders can arise from the skewed salience of competing interests, as reflected by relative levels of neural activity in the corresponding areas of the brain.

A potential architecture for artificial agents

(Kelley, 2005) argues that in their broadest possible sense,

emotions are required for any organism or species to survive. They allow animals to satisfy needs and act more effectively within their environment. If robots are to survive as effectively then they also need equivalent systems.

Implementing these systems and seeing first hand how they benefit artificial agents should help us appreciate the reasons why natural agents are endowed with emotions. It is envisaged that the following architecture should prove useful for the design of artificial agents. The premise behind this is the question; how can we design agents to adapt to unknown environments when we ourselves do not know what those environments will be? The approach taken here is to concern ourselves with the features and resources that are consistent between the different environments that the agent may inhabit. The state of these features shall then be signalled to a self-organising agent controller so that it can adapt accordingly.

Separate subsystems can be hard-coded to look out for certain sensory features and to provide a signal to the neural network controller using modulators. So for example, separate subsystems may be required to recognise the difference between a clean carpet and a dirty one, or the edge of some stairs. A subsystem that recognises a dirty carpet could send out a signal to agitate the neural network controller until the agent started to clean the carpet. Another subsystem could decrease network activity using a neuromodulator when the agent started to recharge batteries that were low. If an agent with such a controller as presented in this paper needs to replenish a critical resource or keep its physiological variables within a critical range, then it will do so by choosing actions that result in lower network activity.

For an agent to maintain homoeostasis using a self-organising system, there must be a non-linear increase in the activity of the network controller in order to signal that equilibrium of its physiological variables needs to be restored. The relative importance of repleting a resource grows as it is depleted. Taking the example of a battery level for a robotic agent, a change from 2% to 1% is absolutely critical, far more say than a change from from 15% to 14%. The increase in network activity needs to be amplified, it needs a gain control.

The signal need not carry much information, but it must be acted upon when the level of a resource becomes critical. It must increase the activity of the parts of the system that allow homoeostasis to be re-asserted. Once this has been achieved, the signal is no longer required and the system can settle back into a stable state. The more significant the increase in network activity, the more the self-organising system is driven to minimise it. As with Panksepp's example of food tasting better the more hungry we are, neuromodulators could work together to repulse and attract the agent towards specific actions or environments. If there is a significant increase in activity because the agent has a particular resource that is critically low, then a neuromodulator that decreases activity in the affected parts of the brain will have a much stronger effect.

The example of homoeostasis was used because the need to maintain equilibrium of physiological processes is con-

stant, regardless of the environment that an agent may inhabit. This mechanism need not be constrained to maintaining homoeostasis though. Significant changes in network activity could be triggered because of specific external, rather than internal, sensory stimuli. An increase in network activity was used in this paper to drive an agent to re-explore the effect of actions that had previously shown to be less than optimal. But it could probably also be used to implement the role of disgust by making certain external stimuli aversive, driving the agent away from it.

Satisfying emotions could help encourage an agent to perform otherwise neutral actions. There are situations whereby an agent may need to perform an action that will indirectly minimise network activity because it leads to a more promising environment. So for example, an agent may have learnt from previous experience that being in a particular environment increases the chances of a specific resource being repleted. Learning to recognise such an environment and inducing satisfying emotions that minimised activity in the network would attract the agent towards it.

An agent may also need to perform an action that does not help in maintaining homoeostasis at all. In the case of an artificial agent developed for a specific purpose, we may wish to bias it to perform actions or behaviours which are of no benefit to the robot but for example, provides us with cleaner carpets.

References

Avila-García, O., & Cañamero, L. (2005). Hormonal modulation of perception in motivation-based action selection architectures. In *Proceedings of the symposium on agents that want and like: Motivational and emotional roots of cognition and action at the aisb-05 conference* (p. 9-16). The society for the study of artificial intelligence and the simulation of behaviour.

Bi, G. Q., & Wang, H. X. (2002). Temporal asymmetry in spike timing-dependent synaptic plasticity. *Physiol Behav.*, *77*(4-5), 551-555.

Blanchard, A., & Cañamero, L. (2006). Developing affect-modulated behaviors: Stability, exploration, exploitation, or imitation? In F. Kaplan et al. (Eds.), *Proc. 6th intl. workshop on epigenetic robotics, lund university cognitive studies, 128*.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. Quill.

D.Beer, R.(1995). A dynamical systems perspective on autonomous agents. *Artificial Intelligence*, *72*, 173 - 215.

Di Paolo, E.(2003). Evolving spike-timing-dependent plasticity for single-trial learning in robots. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, *361*(1811), 2299-2319.

Evans, D.(2002). The search hypothesis of emotion. *British Journal for the Philosophy of Science*, *53*(4), 497-509.

Fellous, J.-M. (1999). The neuromodulatory basis of emotion. *The neuroscientist*, *5*(5), 283-294.

Fellous, J.-M. (2004). From human emotions to robot emotions. In *Architectures for modeling emotions: Cross-disciplinary foundations. papers from the 2004 aaai spring symposium* (p. 37-47). AAAI Press.

Gotlib, I. H., & Hamilton, J. P. (2008). Neuroimaging and depression: Current status and unresolved issues. *Current Directions in Psychological Science*, *17*(2), 159-163.

Hertz, J., Krogh, A., & Palmer, R. G. (1991). *Introduction to the theory of neural computation*. Addison-Wesley Longman Publishing Co., Inc.

Heylighen, F. (2000). The science of self-organization and adaptivity. In *The encylopedia of life support systems* (p. 253-280). EOLSS Publishers.

Karmarkar, U. R., & Buonomano, D. V. (2002). A model of spike-timing dependent plasticity: One or two coincidence detectors. *Journal of Neurophysiology*, *88*, 507-513.

Karmarkar, U. R., Najariana, M. T., & Buonomano, D. V. (2002). Mechanisms and significance of spike-timing dependent synaptic plasticity. *Biological Cybernetics*, *87*, 373-382.

Kauffman, S. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford University Pres.

Kelley, A. E. (2005). Who needs emotions? the brain meets the robot. In (p. 29-77). Oxford University Press.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. A Bradford book. The MIT Press.

Koch, C. (1999). *Biophysics of computation*. Oxford University Press.

LeDoux, J. E. (1998). *The emotional brain*. Simon & Schuster.

Malsburg, C. von der. (2003). The handbook of brain theory and neural networks. In (p. 1002-1005). The MIT Press.

Minsky, M. (1988). *The society of mind*. Simon & Schuster Inc.

Nesse, R. (1990). Evolutionary explanations of emotion. *Human Nature*, *1*(30), 261-289.

Noble, J. (1997). The scientific status of artificial life. In *Proceedings of the fourth european conference on artificial life.*

Panksepp, J. (1998). *Affective neuroscience: the foundations of human and animal emotions*. Oxford University Press, Inc.

Parussel, K. M. (2006). *A bottom-up approach to emulating emotions using neuromodulation in agents*. Unpublished doctoral dissertation, University of Stirling.

Parussel, K. M., & Cañamero, L. (2007). Biasing neural networks towards exploration or exploitation using neuromodulation. In J. M. de Sá, L. A. Alexandre, W. Duch, & D. Mandic (Eds.), *Icann 2007: Proceedings of the 17th international conference on artificial neural networks part ii* (Vol. 4669, p. 889-898). Springer-Verlag.

Parussel, K. M., & Smith, L. S. (2005). Cost minimisation and reward maximisation. a neuromodulating minimal disturbance system using anti-hebbian spike timing-dependent plasticity. In *Proceedings of the symposium on agents that want and like: Motivational and emotional roots of cognition and action at the aisb-05 conference* (p. 98-101). The society for the study of artificial intelligence and the simulation of behaviour.

Rolls, E. T. (1999). *The brain and emotion*. Oxford University Press.

Rolls, E. T. (2005). Who needs emotions? the brain meets the robot. In (p. 117-146). Oxford University Press.

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive hebbian learning through spike-timing-dependent plasticity. *Nature neuroscience*, *3*(9).

Wehmeier, U., Dong, D., Koch, C., & Essen, D. van. (1989). Modeling the mammalian visual system. In C. Koch & I. Segev (Eds.), *Methods in neuronal modeling: From synapses to networks* (p. 335-360). The MIT Press.

Wörgötter, F., & Porr, B. (2004). Temporal sequence learning, prediction and control - a review of different models and their relation to biological mechanisms. *Neural Computation*, *17*, 1-75.

Zhang, Q., & Lee, M. (2009). Analysis of positive and negative emotions in natural scene using brain activity and gist. *Neurocomput.*, *72*(4-6), 1302-1306.