EXPLORING THE NEEDS OF EMOTIONS IN A MULTI-AGENT SYSTEM

Christopher S. Tingley

MEng Computer Science and Cybernetics, siu01cst@rdg.ac.uk

ABSTRACT

Human emotions have fascinated scientists for many years because of the diversity of findings between differing fields of study. Even though there is no clearly defined answer as to what emotions are and why we have them; we can still apply our limited knowledge of emotions to other areas of science. Distributed artificial intelligence; specifically Multi-Agent Systems (MASs), is a relatively new area of science that has only come about within the last 10 years. As a result of its' infancy, many MASs suffer from a number of common fundamental problems. Planning, scheduling and goal management are all inter-related problematic areas of MASs. These are caused by large search domains, multiple dimensionality, non-linearity, conflicting interests and NP-complete problems, all of which are seen to be echoed, but more importantly solved in human systems. This paper will show how human systems overcome these problems using simple emotional mechanisms. A hypothesis will be made that many challenges within MASs can be solved by implementing mechanisms which mimic the purpose of human emotions.

1. INTRODUCTION

The brain is largely and unknown entity and is popularly (but incorrectly) stated that only 10% is utilised. Realistically this figure is nearer 100% but does perhaps represent the amount of the brain that is understood. In recent years much research has given us a greater understanding of what it does, how it works and why.

It is important to note that the split between research into the physical, physiological and psychological functions and processes of the brain has equal weighting between the disciplines. It is easy to see the appeal of studying human emotions and moods since these three independent disciplines have supplied three differing opinions as to what emotions are and what role they play within each discipline. This research will make up the first half of this survey paper while the second half will address the differing architectures in single and multi-agent systems.

This paper will cover important findings relating to human emotions and their possible novel applications within MASs. Ultimately, this knowledge will be used in order to create emotional agents as part of an emotional MAS, with the goal being to improve upon the more classical approaches to the architecture design behind current MASs. Other survey papers on similar subject matter include [7] and [30].

2. EMOTIONS AND MOODS

2.1. What are emotions?

Emotions are known to be linked into how humans make decisions and help guide the way the species lives in both work and play. They are important in the ways in which humans communicate and how humans express themselves both verbally and non-verbally. For these reasons, emotions command an arsenal of power opening up possibilities to use them in ways which may not have been thought of before.

So, what are emotions? This is a difficult question due to the differing conclusions which can be drawn from vastly different research areas; and as a result, one to which there is no clearly defined answer. Emotions by their nature are subjective and therefore non-quantifiable. Although there has been much research in this field (Broca [4], Darwin [8], James [18], MacLean [21], Papez [26] and Pavlov [27] to name a few) there has been no commonly agreed definition. From [22] §1-3, Minsky lists popular definitions of emotions from a selection of dictionaries:

- The subjective experience of a strong feeling.
- A state of mental agitation or disturbance.
- A mental reaction involving the state of ones body.
- A subjective rather than conscious affection.
- The part of consciousness that involves feeling.
- A non-rational aspect of reasoning.

As seen by this varied list of meanings, a solid conclusion can not be drawn as to what emotions are and how or why they are created. Minsky points out that “if you didn’t yet know what emotions are, you certainly wouldn’t learn much from this”. The word “emotion” in language is used to describe a massive range of differing mental states, all of which are linked and cross-referenced against a huge array of other human-like subjective “feelings”.

Affective Neuroscience has linked emotions to a group of structures in the center of the brain known as the limbic system which consists of the cingulate cortex, hippocampi and the hypothalamus among others. In recent years other non-limbic structures have also been linked to emotion [36] and are found throughout the brain, these include:

- **Anterior cingulate** - primarily involved with attention and motivated behaviour.
- **Amygdala** - thought to detect and learn which parts of our surroundings are important to our mental state and how they may have emotional significance. Also believed to be involved with the generation of negative emotions.
- **Insula** - regulates the body’s automic functions such as breathing and heart rate etc. Also known to processes taste information and thought to be connected with the invocation of disgust.
- **Prefrontal cortex** - thought to compute outcomes due to our actions.
- **Ventral striatum** - plays a large part in goal-orientated (positive) emotions. Known to demonstrate increased activity when encountering addiction.

By the 1980’s, the reward centre of the brain had been mapped out, as well as the discovery of dopamine - the chemical transmitter which gives us pleasure. This “dopamine system” was linked to “rewarding animals for doing things with survival value - eating or having sex for example” [28].

Another basic emotion which has been the centre of much research is fear. [16] states: “without a proper fear motivation, animals would not be capable of taking the proper safeguards to avoid harm to themselves or others from simple and routine actions, such as walking around a lege or crossing a busy street”. Research into this area has told us that sensory input arriving at the amygdala drives most fear responses [3], but other Higher Order Processes (HOPs) and even memories can also trigger fear responses [19] [20].

A conclusion may be drawn that emotions are simply a result of neuro-mechanisms linked to actions and consequences, which might explain the physical processes of emotions and why humans have them. One particular text of note that shares this hypothesis is “The Brain and Emotion” by Rolls [29]. Rolls defines emotions as “the states elicited by reward and punishers, including changes in those rewards and punishments”, by which “a reward is anything for which an animal will work, i.e. a Positive Reinforcer” and “a punishment is anything that an animal will work to escape or avoid, i.e. a Negative Reinforcer”. This description is able to provide a straight-forward explanation of the role of emotions within animals. It is considered that emotions are a result of the actions humans take and are a sum of the rewards and punishments obtained along the way - much like the actions and consequences that result from chemical processes in the brain.

Rolls’ definitions of positive and negative reinforcers are similar to the principles beind Pavlov’s classical conditioning [27]. Classical conditioning is the process of pairing two stimuli to cause the learning of the second stimulus in order to recreate the response associated with the first stimulus. As young children, humans may associate emotional feelings of anxiety against punitive figures which become “conditioned” and continue into adult life. The same process can be used in clinical settings in later life to cure phobias and other such psychological phenomenon by unlearning previously conditioned responses.

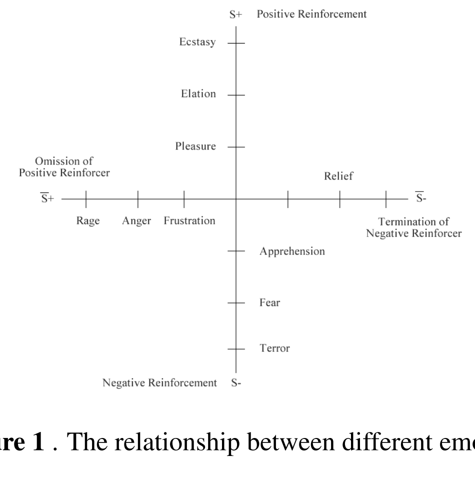

Figure 1 . The relationship between different emotions

Rolls provides Figure 1, pictorially demonstrating the links between differing emotions. The intensity of emotion increases from the origin of the graph where the axes represent one of the following:

1. **Positive reinforcer** - leading to pleasure or;
2. **Negative reinforcer** - leading to fear or;
3. **Omission of a positive** - reinforcer leading to anger or;
4. **Termination of a negative reinforcer** - leading to relief.

Rolls also provides a number of reasons for humans having emotions. The discussion will illustrate the similarities between this list and the the hypothesised reasons for adding emotions to agents.

1. **Elicitation of autonomous responses** - some responses are not simply taxis responses (chemically excitative). For example the fight or flight response in animals is based upon an emotional response.
2. **Flexibility of behavioral responses** - for example, an animal may learn to do something by either (1) obtaining a reward or (2) avoiding a punishment. Either method can arrive at the same mood state.
3. **Motivation** - if an animal has learnt that an action has obtained a reward in the past, then the animal will attempt the action again in order to receive the reward. The opposite of avoiding an action to avoid punishment.
4. **Communication** - both verbal and non verbal. It is possible for humans to convey more information by communicating our emotions along side our spoken word.
5. **Social bonding** - possibly on a non-communicative level.
6. **Generalisation** - i.e. being able to group situations which may use the same method to obtain a reward (or avoid punishment) in an effort to reduce the amount of situations we need to learn. This is also useful when humans are put in new situations where the the outcome is not known. By comparing this with similar situations, humans are able to make intelligent judgements.
7. **Mood state can affect cognitive evaluation of an event or memory** - are memories stored differently when our mood is confused?
8. **Aiding storage of memories** - are we able to store memories based on our emotional state?
9. **Recalling memories** - connected to (7) and (8).
10. **Persistent motivation by enduring emotion** - e.g. if an action produces a pleasant emotion, it is desirable to continue the action. This is subtly different from (3) which encourages us to start an action in the first place as opposed to continuing an action we have started.

Finally, Rolls puts forward four levels of complexity within brain design where it is important to note that only humans exhibit all four levels of complexity.

1. **Simple taxis responses** - very similar to the locomotion - tropism effect in plants.
2. **Classical conditioning** - see page 3.
3. **Two process learning**:
   (a) Classical conditioning followed by the advancement of...
   (b) Attempting to find a the stimulus in order to receive a reward
4. **Multi-step one off planning** - e.g. learning to find a stimulus when already confident of the reward (similar to 3b).

2.2. What are moods?

Moods are often linked to emotions, but suffer from the same problem in definition. "A mood is an animal's quantifiable affective state which is comprised of a number of emotions" [37]. This is unlike the definition forwarded by Rolls who states that moods can: "happen without external sensory input". This suggests that moods may consist of components other than emotions and that humans can control their mood without needing an external stimulus to generate an emotion leading to a mood.

As a side note, the word mood is also used within the english language to mean a range of different things. In music for example, mood is described as the feeling created by a work of music. In art, mood is used to describe the expression of emotion within a painting or other piece of artwork. As with emotions, the exact meaning of the word is difficult to define and again, may simply be a non-quantifyable subjective measure.

3. AGENTS AND ARCHITECTURES

3.1. Historical architectures

An agent architecture can be defined by a collection of components or modules (layers) with clearly defined interfaces and connections, structured in a hierarchial manner. Components can be descried on a number of levels. Low-level components such as actuators, sensors or circuit boards (on a physical agent i.e. a robot) and the connections between them. Mid-level components may refer to the programmatic behaviours of the agent through observation of the agents' interaction with the world via its actuators and sensors (whether they are simulated or real-life). High-level encompasses the emergent behaviours of multi-agent systems or meta-management layers of single agent systems (c.f. Moving towards emotional architectures). Agent architectures must have some degree of modularity allowing components to be easily changed, removed, appended, swapped or modified without having to reprogram or redesigning the remaining or underlying architecture.

Classical Artificial Intelligence (AI) in mobile robotics started with purely reactive and/or deliberative systems [23]. These were effective at dealing with immediate situations (i.e. reacting on "instinct") by matching situations to actions through a linear process see Figure 2 (n.b. reactive systems did not plan or model). Reactive systems react immediately to inputs akin to a reflex reaction in an animal. Deliberative systems use basic goal management to compare inputs against predefined goals and take the action that is best suited to achieve the end goal.

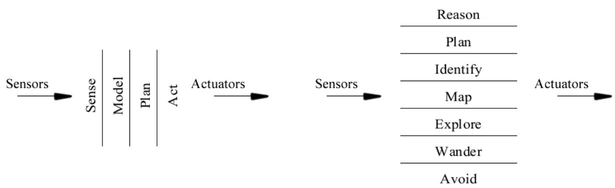

Figure 2 . Deliberative vs Layered Architectures

A more useful architecture is one comprised of a number of layers see Figure 2. To this extent, robot (and other agent-based) architectures were given a new lease of life when Brooks published his seminal paper entitled "A robust layered control system for a mobile robot" [5]. This introduced the concept of having multi-layered subsumption based architectures, see Figure 3, resulting in a powerful programming paradigm allowing complex interactions and behaviours to emerge. The subsumption architecture is still used very much today, albeit with minor modifications between incarnations.

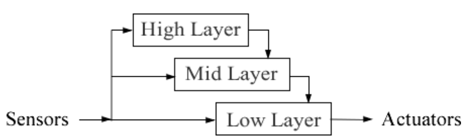

Figure 3 . The Subsumption Architecture

A third and final "hybrid" architecture combines (1) the basic reactive and deliberative control mechanisms along with (2) a subsumption based layering structure and (3) a symbolic or other higher level process which overlooks the goals of the system. This form of architecture provides a number of advantages over other forms of architecture. These are listed below:

- A robust control system.
- Varying levels of competence.
- Layers of control.
- Parallel computation abilities.
- Simple behavioral programming.

3.2. Moving towards emotional architectures

Sloman [33] put forward a "Triple Layer Perspective" (hybrid) architecture produced from a low-level reactive layer, a middle deliberative layer and a higher "Meta-management" layer. This architecture allows the classical reactive and deliberative layers to operate as ever, but with the addition of a meta-management layer designed to allow other operations such as emotions to be placed within the system.

Sloman suggests that "the absence of meta-management was a major factor in the fragility and incompetence of many old AI systems (e.g. they could not tell when the were reasoning in circles, or solving a minor variant of a previously solved problem)". In essence they would never be able to exhibit the complex abilities of an animal system. By adding a meta-management level it is hypothesised that the following abilities could be realised:

- Ability to monitor, categorise, evaluate and (to some extent) control other internal processes.
- Control attention and thought processes - for example, a deliberative layer is made up from a series of "what ifs?" generated by the inputs into the system. Some of these "what ifs?" may not be relevant to the current task and so can be ignored, thus directing attention to the most important considerations.
- Varying levels of sophistication e.g. having to depend on social learning.
- Self evaluation - e.g. How do I feel?
- Characterise mental states - e.g. What would I call how I feel (happy, frustrated, etc.)?

In essence, adding a meta-management layer is very similar to the hybrid architecture described in 3.1. Assuming this is possible, what would emotions in this meta-management layer actually do? [31] gives 12 roles of emotion within emotional agents:

- **Action selection** - what to do next based on current emotional state(s) or mood.
- **Adaption** - short or long-term changes in behaviour due to emotional state(s).
- **Social regulation** - communicating or exchanging information with other agents via emotional expressions.
- **Sensory integration** - emotional filtering of data or blocking of specific sensory integration.
- **Alarm mechanisms** - fast, reflex-like reactions in critical situations that interrupt other processes (c.f. reactive control).
- **Motivation** - creating motives as part of an emotion coping mechanism.
- **Goal management** - creation of new goals or reprioritisation of existing ones (the likely subject of the authors advancement within emotional agent research).
- **Learning** - emotional evaluations as Q-values in reinforcement learning.
- **Attentional focus** - selection of data to be processed based on emotional evaluation
- **Strategic processing** - selection of different search strategies based on overall emotional state (again, another area which may be researched further by the author).
- **Self model** - emotions as representations of "what a situation is like for the agent", e.g. self evaluation and or categorisation.

3.3. An introduction to Multi-Agent Systems

Sycara [35] provides an excellent introduction to MASs. Sycara identifies the characteristics of MASs as: (1) each agent has incomplete information or capabilities for solving the problem and, thus, has a limited viewpoint; (2) there is no global system control; (3) data is decentralised and (4) computation is asynchronous. The most important uses of MASs are determined:

- To solve problems that are too large for a centralised agent.
- To allow for the interconnection and interoperation of multiple existing (homogeneous or heterogeneous) systems.
- To provide solutions to a problem that can naturally regarded as a society of autonomous interacting agents.
- To provide solutions that efficiently use information sources that are spatially distributed.
- To provide solutions in situations where expertise is distributed.
- To enhance performance along the the dimensions of: (1) computational efficiency, (2) reliability, (3) extensibility, (4) robustness, (5) maintainability, (6) responsiveness, (7) flexibility and (8) reuse.

4. DISCUSSION

It is hypothesised that a multi-agent systems can be created that are capable of controlling their operations based upon their emotions and moods, therefore avoiding common problems associated with MASs. This discussion will support this hypothesis by calling upon the research introduced in sections 1 to 3.

As seen from Rolls' uses of emotions within humans and Scheutz's proposed uses of emotions within agent-based systems, many of the traits can be matched. This is a positive indication that emotional agents would produce promising results.

Moods have not been shown to have the same importance in humans as emotions have. As a result of this, it is considered that moods pose no use to MASs. Another area (which would be the subject of a separate paper) is one of learning within MASs. It is clear that humans have capabilities to learn, providing the ability to solve complex problems and avoid issues associated with agent-based systems. The ability to learn was touched upon in section 1.1 with Rolls' four levels of brain complexity.

The uses of MASs make them ideal for a number of situations. For this reason MASs provide the ideal platform for emotional agent research. This is due to a few important reasons; primarily, a MAS presents a suitably complex system. This allows intricate interactions to be observed between agents with the possibility of emergent behaviours either (a) evolving from a system capable of learning; or (b) through detailed pre-programed strategies.

Some improvements of MASs using human inspired emotions are suggested:

- **Task allocation** - the study of breaking down large tasks into smaller ones and then distributing them amongst a number of agents. This is an essential part of any MAS.
- **Task monitoring and scheduling** - the ability to reallocate tasks within a system if an agent fails unexpectedly or other agents finish tasks before or behind schedule, i.e. by allowing agents to experience motivation and frustration.
- **Distributed goal management** - giving each agent in the system the ability to modify its task based upon the agents' emotional state.

The authors' research will implement an emotional MAS architecture in order to demonstrate that an emotional MAS can preform better at a given task. One of the suggested improvements given above can be used as a benchmark to compare emotional with non-emotional MASs.

5. CONCLUSION

Emotions play a very important role within human systems and these emotions guide, if not control human behaviours. Without these emotions, human systems would encounter problems that would be difficult to circumvent using rational mechanisms.

When looking at the problems associated with MASs, it has been shown that many of them are similar to the problems that human systems avoid by using emotional control mechanisms. By using human-like emotions as a non-rational control mechanism within a MAS, it is possible to avoid some of the pitfalls that have been identified. It is a common trait amongst animals, especially humans that they can work together in groups; akin to a MAS. This fact strengthens this hypothesis.

The research conducted to form the basis of this paper has produced four major findings. (1) The subsumption architecture is a solid foundation for any new emotion-based architecture designed for use in a MAS and should be used as such. (2) Current findings suggest moods can be generated without emotions (and or sensory input) and should therefore not be used within MAS. Implementing moods would require generation using internal random selection which is unlikely to produce desired behaviour. (3) All animals have reactive mechanisms, many also have deliberative mechanisms. We should therefore expect a MAS to make use of these processes. Following from that, only a few species have meta-management abilities. Finally, (4) in order for a MAS to even attempt to be as sophisticated as any of these systems, meta-management mechanisms should be implemented within them.

Acknowledgements: Thanks go to Dr. William Browne for his invaluable help with collating the research material for this paper and providing useful and challenging discussions. Produced using LaTeX.

6. REFERENCES AND BIBLIOGRAPHY

[1] Aylett, A. and Delgado, C. Emotion and agent interaction. Technical report, CVE, University of Salford, Salford, England, 2001.

[2] Bond, M. Science of happyness, part 1: 10 keys to a contented life. New Scientist, pages 40–47, October 2003.

[3] Bradley, M. M. and Lang, P. J. Measuring emotion: Behaviour, feeling and physiology. In Lane, and Nadel, M. S, editors, Cognitive neuroscience of emotion, pages 242–276, NY, USA, 2000. Oxford Press.

[4] Broca, J. Anatomie comparée des circonvolutions cérébrales: le grande lobe limbique. Anthropol, 1:385–498, 1878.

[5] Brooks, R. A. A robust layered control system for a mobile robot. Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA, 1985.

[6] Brooks, R. A. Intelligence without reason. In Mylopoulos, John and Reiter, Ray, editors, Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91), pages 569–595, Sydney, Australia, 1991. Morgan Kaufmann publishers Inc.: San Mateo, CA, USA.

[7] Canamero, D. Issues in the design of emotional agents. In Canamero, D., editor, Emotional and Intelligent: The Tangled Knot of Cognition: Papers from the 1998 Fall Symposium, pages 23–27, 1998.

[8] Darwin, C. The Descent of Man, and selection in relation to sex. John Murray, London, 1871. Reprinted in 1981 by Princeton University Press.

[9] Davis, D. N. Computational emergence and computational emotion. Technical report, Department of Computer Science, The University of Hull, Hull, England, 2000.

[10] Davis, D. N. Modelling emotion in computational agents. Technical report, Department of Computer Science, The University of Hull, Hull, England, 2000.

[11] Davis, D. N. Cellular automata, computational autonomy and emotion. Technical report, Department of Computer Science, The University of Hull, Hull, England, 2001.

[12] Davis, D. N. Multiple level representation of emotion in computational agents. Technical report, Department of Computer Science, The University of Hull, Hull, England, 2001.

[13] El-Nasir, M. S., Loerger, T. R., and Yen, J. Peteei: A pet with evolving emotional intelligence. In International Conference on Autonomous Agents, pages 9–15, 1999.

[14] Gadanho, S. and J., Hallam. Emotiontriggered learning for autonomous robots, 1998.

[15] Gage, A., Murphy, R., Valavanis, K., and Long, M. Affective task allocation for distributed multi-robot teams, 2004.

[16] Hannon, C. J. Emotion-based control mechanisms for agent systems. In International Conference on Information Systems and Engineering, Fort Wort, TX, USA, June 2003. Texas Christian University.

[17] Hertz, J. Computing with attractors. In The handbook of brain theory and neural network, pages 230–234, MA, USA, 1998. MIT Press, Cambridge.

[18] James, P. The Principles of Psychology. Unknown, 1890.

[19] Ledroux, J. E. In search of an emotional system in the brain: Leaping from fear to emotion and consciousness. In Gazzaniga, M. S., editor, The Cognitive Neurosciences, pages 1049–1061, Cambridge, MA, USA, 1995. MIT Press.

[20] Ledroux, J. E. The emotional brain. Simon and Schuster, NY, USA, 1996.

[21] Maclean, P. D. Some psychiatric implications of physiological studies on frontotemporal portion of limbic system (visceral brain). Electroencephalogr Clin Neurophysiol Suppl. 4, 4:407–418, 1952.

[22] Minsky, M. The emotional machine. Unpublished, 2003.

[23] Moravec, H. P. The stanford cart and the cmu rover. In Proc. IEEE, volume 71, pages 872–884, July 1983.

[24] Murphy, R., Lisetti, C., Irish, L., Tardif, R., and Gage, A. Emotion-based control of cooperating heterogeneous mobile robots. In IEEE Transactions on Robotics and Automation: special issue on Multirobots Systems, 2001.

[25] Napel, G. Modelling evolution using asynchronous stochastic cellular autonoma. Technical report, CVE, University of Salford, Salford, England, July 2003.

[26] Papez, J. W. A proposed mechanism of emotion. J Neuropsychiatry Clin Neurosci. 7, 1:103–112, 1937.

[27] Pavlov, I. P. Conditioned reflexes. Routledge and Kegan Paul, London, 1927.

[28] Phillips, H. Science of happyness, part 2: What is pleasure for? New Scientist, pages 36–43, October 2003.

[29] Rolls, E. T. The brain and emotion. Oxford University Press, 1999.

[30] Scherer, K. R. Criteria for emotion-antecedent appraisal: A review. In Hamilton, V., Bower, G., and Frijda, N., editors, Cognitive perspective on Emotion and Motivation, pages 89–126. Kluwer Academic, 1988.

[31] Scheutz, M. Useful roles of emotions in artificial agents: A case study from artificial life. In AAAI, pages 42–48, 2004.

[32] Sengers, P., Kulessa, T., and Geng, W-D. The influencing machine - emotions in interation. Technical report, Media Arts Research Studies, GMD.IMK, Sankt Augustin, Germany, 2000.

[33] Sloman, A. When will real robots be as clever as in the movies? In ASE Conference, Birmingham, UK, January 2003. Department of Applied Mathematics and Theoretical Physics, University of Birmingham.

[34] Sloman, A. and B., Logan. Cognition and affect: Architectures and tools. In Sycara, Katia P. and Wooldridge, Michael, editors, Proceedings of the 2nd International Conference on Autonomous Agents, pages 471–472, New York, 9-13, 1998. ACM Press.

[35] Sycara, K. Multiagent systems. AI Magazine, 10(2):79–93, 1998.

[36] Wikipedia, . Affective neuroscience, October 2005. http://en.wikipedia.org/wiki/Affective_neuroscience.

[37] Wikipedia, . Emotional moods, October 2005. http://en.wikipedia.org/wiki/Emotional_mood.

[38] Wright, I. Reinforcement learning and animat emotions. Technical Report CSRP-96-4, Unknown, January 1996.

[39] Wright, I. Emotional Agents. PhD thesis, University of Birmingham., 1997.