Systemic Design: A Methodology For Investigating Emotional Phenomena

(In the process of being updated)

Tim Read*
Cognitive Science Research Centre
School of Computer Science
University of Birmingham
Birmingham, UK
Email: T.M.Read@cs.bham.ac.uk

August 28, 1993

*This work is funded by the S.E.R.C., grant number: 91313564, and will be presented as a poster at the 1993 Australian Joint Conference on Artificial Intelligence, Melbourne, 16-19th of November.

Abstract

In this paper I introduce Systemic Design as a methodology for studying complex phenomena like those commonly referred to as being emotional. This methodology is an extension of the design-based approach to include: organismic phylogenetic considerations, a holistic design strategy, and a consideration of resource limitations. It provides a powerful technique for generating theoretical models of the mechanisms underpinning emotional phenomena, the current terminology associated with which is often muddled and inconsistent. This approach enables concepts and mechanisms to be clearly specified and communicated to other researchers in related fields.

1 Introduction

Obtaining a theoretical understanding of emotions is not an easy task partly because researchers disagree on the definition of “emotion”. We must therefore define the specific emotional phenomenon in question before trying to postulate its formation and functionality. This should emphasise the composition of the phenomenon in question rather than terminology commonly associated with it (e.g. an ‘angry state’), in order to reduce the chance of falling into terminological pitfalls which hinder common agreement (Read and Sloman, 1993).

The benefits of using a design-based approach for developing a theoretical understanding of such phenomena has been put forward by Sloman (1993b). It is a powerful tool for such a task, but can be criticised in several ways. Firstly it can be argued that an agent’s architectural functionality exists as a network of interdependencies which can’t be split into unitary phenomena to be modelled; i.e. the operation of each mechanism within the architecture is to some extent dependant on the other mechanisms present. Furthermore trying to model all the phenomena together may not be possible due to the inherent complexity. Secondly, systems built using the design-based approach to fulfil a set of biologically detached abstract requirements will not generate any useful theoretical knowledge about human capabilities as the requirements are not under the same constraints that apply to naturally evolved intelligent systems (Clark, 1989).

Lastly, Pfeifer (1993) raises another doubt by suggesting that as there is no clear understanding of the relation between mechanism and behaviour, it is impossible to build accurate models of the underlying mechanisms. The result of the argument Pfeifer (1993) puts forward is to suggest that an understanding will emerge by building simple agents first, and then working toward more complex agents. Each subsequent agent is designed with the understanding gained from the way in which the mechanisms worked in previous ones. This technique is unlikely to succeed for a couple of reasons: Firstly there are too many intermediate agents to design and build between very simple life forms and intelligent agents like man. Secondly any design decision made for one agent may well limit how future agents can function, assuming they inherit variations of the mechanisms tested in earlier agents: which is the basis of the approach. In this paper I wish to introduce a methodology called Systemic Design, which is an advancement on the design-based approach and addresses the problems levelled against it here.

2 Systemic Design

Systemic Design is composed of four things: The design-based approach, organismic phylogenetic considerations, a holistic design strategy, and a consideration of resource limitations.

2.1 The Design-based Approach

This approach to solving a problem is based around performing a requirements analysis of the problem in hand, followed by an exploration of the possible mechanisms to meet these requirements (the exploration helps to clarify the requirements). Finally a design is produced for mechanisms that best meet all the requirements. A key fact about this technique is that there is not a unique way of solving any problem.

2.2 Organismic Phylogenetic Considerations

The role of phylogenetic considerations in the functionality of an organism is discussed by Dawkins (1986). It is also placed in the framework of computer models of live organisms by Clark (1989). The essence of the argument is that in order to understand the functionality of a currently existing organism, it is necessary to consider the nature of the problems encountered (and solutions found) during the evolution of its species. It may well be the case that a problem that caused some feature of an organism's architecture to come into existence is no longer around. Hence the design-based approach, used to solve the organism's current problem set, will not produce the same type of architecture as would a design strategy that incorporates phylogenetic considerations.

When considering the role of evolution via natural selection on the current architecture of an intelligent agent, it is tempting to focus on species development at a course level (i.e. modelling the evolution of populations whose abilities and interactions are guided using simple rules), rather than on the nature of the previous problems the agent's ancestors will have solved at particular points in the past. Before adopting such a focus, it is important to be clear what the goals of incorporating phylogenetic considerations are here. The principal goal is not to build a model of every step in the evolutionary process, but to generate a theory of the mechanisms that under-pin human functionality. A related point is the difficulty in using a study of species evolution, to progress from a simple organism to human capabilities. As Dawkins (1986) points out, life as it is now is one solution to the problems of existence, and there is no reason to assume that a process starting again would end up at the same point in evolutional space. For example, if there are only ten paramaters that define the evolutional space of all life forms (i.e. ten degrees of freedom), then in ten steps through that space, there are 10^{10} permutations of agent available; i.e. 10,000,000,000 combinations! Hence the chances of following the exact path between two points in the evolutional space (from a very simple organism to a complex agent like man, in many million steps) is very small!

My argument about the use of phylogenetic considerations concerns the importance of understanding the types of problems that our ancestors would have had to solve, and how the solutions that they evolved might have influenced the mechanisms we have today. This does not require a very detailed understanding of every evolutionary step. Instead it requires judicious consideration of key types of problems and their solutions, in the context of the likely levels of architectural sophistication that would have been present at the time. For example if we are interested in the mechanisms underpinning an apparently function less human phenomenon X, we can consider the types of problems and situations our ancestors might have experienced that might have required X. When considered from a different time phenomenon X may be of benefit, or even essential to survival. If we find some need for a capability Y which is in some sense similar to X, but possibly simpler, and there is some empirical evidence to support its presence in our ancestors, it will be informative to consider the mechanisms responsible for Y as being present in some sense in humans that exhibit X. If the theoretically postulated mechanisms preclude some other capability or phenomenon that we exhibit, they can be reworked or abandoned and different phylogenetic considerations taken into account which fits all the capabilities.

2.3 A Holistic Design Strategy

The design-based approach is a very powerful tool in working toward an understanding of the functionality of an unknown system. A lot of design carried out in the fields of computer science and engineering is primarily top down; i.e. models are built by formulating high level requirements of performance, and then decomposed down to form the specification for functional mechanisms that make up the system. A lot of 'post hoc' design done in the fields of psychology, psychobiology, neuropsychology etc is bottom up; i.e. models are built in a data driven fashion, by starting with the low level mechanisms and working up. Both of these approaches have their place in the research field, but a better technique can be made by combining them both. This offers the benefits of each technique in testing and validating the input from the other.

For example if data from brain studies lead to the generation of differing accounts of the role of a particular brain structure in a particular phenomenon, the conflict may be able to be resolved (to some degree) by trying to incorporate both low level accounts into a high level design of some functionality of the organism. This may show that one account is wrong, as it directly conflicts with some other capability that the system provides. Similarly if a high level design reaches a point where there are several different solutions to a particular problem, low level data can be used to decide which is correct. By pulling together data taken from different fields of expertise the scope for mistakes in the generation of theoretical models is limited. This combined top down and bottom up mutual support technique can be called a holistic design strategy, as it places emphasis on the importance of studying a system from as many directions as possible.

2.4 Consideration of Resource Limitations

A key feature of the way a system operates, is the nature of the resources open to it. The degree to which the design-based approach takes account of this fact when modelling human phenomena depends on the nature of resources the designer wishes to place in their model. This in turn will depend on the goals they are pursuing. I would argue that where possible resource limitations should always be taken into account. However, due to the complexity of a system such as a human being it may not be possible to include all the resource limitations in a model, so some degree of selection is required. Also as the computer models can't (yet) be made of 'flesh and blood' (or at least containing the same operational characteristics of a live organism), with their properties and limitations, some compromise must be made with regard to the nature of the limitations incorporated into the model. What follows is a list of the main types of resource limitation that need to be considered when using Systemic Design.

2.4.1 Environment

An intelligent agent has to interact with the world around it in order to survive, so it is important to consider the effects of the world upon its architectural functionality. The effects can be split into two different types: immediate and evolutional. The immediate effects concern the problems and affordances of the moment. Typical features of an environment will be its asynchronous uncertain nature, and the structure and dynamics of the things that occupy it. A problem that has to be overcome by someone trying to build a model world with a situated agent in it, is the degree of simplicity and complexity of the world. If the world is too simple there will be no need for the agent to have rich and sophisticated mechanisms to deal with it.

At first glance it might appear necessary to move away from studying individual agent architectures in order to study evolutionary effects. The pressures of evolution via natural selection on an agent happen at a group or species level within an enviroment. Hence it would seem reasonable to consider how the environment causes different species to compete for available resources (such as food and shelter). The emphasis of the study being placed on the long term developmental effects of interactions within and between groups of agents and other life forms. However the general combinatorial problem with this type of approach was mentioned earlier in this paper (related to the number of degrees of freedom and the number of steps taken through the evolutional space). The solution to the combinatorial problem is to keep the focus of the design exercise (as a way of generating a good explanatory theory of the sorts of mechanisms we have) on the possible mechanisms that subsume the sorts of phenomena that we exhibit.

An example of an evolutionary environmental pressure would be the competition between agents for food, shelter and mates. This is evident in predator prey relationships, sometimes referred to as the 'arms race'. There is some doubt about the use of this analogy (Abrams, 1986) in describing the relationship, however the general principles still provide an interesting example of past environmental problems that had to be solved by our ancestors. If two species are competing (either directly or indirectly) for some resource within the environment, it is informative to consider how their current specialisation would have enabled them to adapt to that pressure. If an example can be taken from the biological literature of such a relationship where one type of creature is known to have won, theories can be generated as to the composition of the winner's antecedent capabilities. Evidence and theory so generated may be able to be transferred into considerations of why we have certain mechanisms and not others. Such justification would remove the need to go further and further back into our evolutionary past to demonstrate why we have a certain apparently useless capability.

2.4.2 Physiology

The role of an agent's physiology in influencing the functionality of its mental architecture can be illustrated by considering the difference between a human being and a futuristic robot (designed to be physically similar at a gross level of description). There are potentially many differences between the physiology of both systems, but for now lets just concentrate on those associated with motion. The robot is defined to act similarly to a human, such that in an uncertain situation, presented with a source of threat, both agents will attempt to run away. A scenario might be something like: both agents are standing looking at an object when a threat appears (i.e. with no prior warning), both run. In this situation the robot will get away first. Why?

Regardless of any benefits the robot might have because it can process information faster than the human, the real advantage (in this example) it has is in the time it takes it to activate its body to move, and its efficiency when moving. The human moves by virtue of contractions of groups of muscles in its legs. These take time to activate and accelerate the human to maximum velocity. The robot on the other hand will have motors moving its legs via servo controlled cables and joints, which can be activated in micro seconds, and propel it to maximum speed in a very short time. The mental architecture of the human has evolved to make the best use of the body in surviving situations like the above; i.e. overcoming the biomechanical and biochemical limitations of its body. It tries to prepare the body for action in several ways, one of which is by raising the level of physiological arousal.

Considering physiological effects like the above one will help to generate requirements for a controlling architecture as present in human beings and other animals. Concentrating the design of a robot purely on the limitations associated with its material make up will not place the same requirements on its controlling architecture as those of a biological life form do. Typical physiological phenomena that need to be accounted for are suggested here. Physiological arousal was mentioned above but is itself a term of contention. It is often referenced in the literature as being a unidimensional concept, which it isn't. Arousal is an ambiguous term. It embraces three response systems (Frijda, 1986) - autonomic arousal, electrocortical changes, and behaviourial activation. Frijda (1986) lists four types of arousal: autonomic, attentional, behaviourial, and electrocortical. He also notes the empirical data to support the claim that different 'emotional states' do not necessarily have different types of physiological arousal.

In situations where there is some potential threat, the body will produce what can be called an 'emergency reaction'. Some key points about this reaction are:

1. Blood is channelled to the head from the extremities to aid processing (Frijda, 1986), (cf the channelling of blood to the stomach after eating to aid digestion, which often produces sleep).

2. Sympathetic nervous system acts in conjunction with adrenaline secretion from the adrena medulla to produce body readiness for 'fight or flight' type behaviours (Gray, 1987).

3. This reaction is present in anxiety states and panic attacks (Clark, 1986).

4. This reaction mobilises energy prior to physical ability. This takes time and must be done in advance if possible. Once prepared, the time taken to perform an action is greatly reduced.

5. Miscellaneous effects include (Frijda, 1986): The heart rate is increased, both rate and strength, to pump more oxygen around the body. The spleen contracts to release stored red blood cells to carry the oxygen around the body. The amount of stored sugar in the liver is reduced for muscle use. Blood pressure increases. Respiration rate increases. Sweating occurs (postulated to aid skin resilience to damage and the ability of the hands to grip). Muscle tension increases ready for action.

6. This reaction is reflex like in its activation, but is also dependent on learning and voluntary control.

An interesting question about the above physiological effects is: If it is of value for the agent to have the properties of the emergency response, why doesn't the agent stay in it for the majority of the time? One obvious answer would be that the energy consumption would be very high, but presumably it could evolve to carry larger supplies of energy. The main reason that this reaction has to be switched off is to facilitate growth, reproduction, and to allow the body to fight infection and prevent exhaustion. All of which require time and bodily resource. Perhaps it is just beyond the limitations of the biochemistry of life for an organism to evolve where these processes can go on whilst in a permanent emergency reaction. The above represents the main sorts of physiological phenomena that need to be related to a mental architecture. Some others are covered in the next section on attentional phenomena.

2.4.3 Attentional Phenomena

The term "attention" is used ambiguously in the literature, and has no generally agreed meaning (Gazzinga, 1975). A promising avenue comes from the notion of control of selection of information (Sloman, 1993a). There are three senses of the term that predominate in the literature (Gazzinga, 1975): Alertness and general receptivity to the input of information, selection of some information from available sources to process, and the degree of 'conscious' effort involved. Each of these points may very well have more than one interrelated mechanism underlying it. Other phenomena put under the banner of attention are (Gazzinga, 1975): diurnal rhythm effects (the change of performance of alertness tasks with body temperature), vigilance tasks (performance drops over time of attending to a task), and phasic alertness (warning of incoming stimulus causes changes in autonomic arousal as the event get close). Also there are the orienting responses where a group of behaviours and physiological reactions become habituated after a few exposures. Alertness to information is also influenced by arousal, so the previous points about arousal not being a unidimensional concept may be important here.

2.4.4 Control and Representation

The nature of the underlying control and representation present in an architecture will dictate its capabilities. The human being uses many different types of control in its functionality. Some go purely through the brain, some go through the brain and the world (via perceptual sensors), and others go through the brain and body. Examples of the types of control include such things as homeostasis (stable control), moving a hand to catch a thrown object (ballistic control), and moving body parts during walking (dynamic (or unstable) control).

Another classification of the types of control that are needed in a human being is between point to point (PTP) and global control. PTP control refers to the dissemination of control information along the nerve fibres that contact the muscle groups to the central nervous system. As I am typing in this paper my brain is sending specific groups of control signals to the muscles in my hands to guide my fingers over the keyboard. Global control can be illustrated by the effects of hormones. When a hormone is released its effects are experienced by several different systems at once, causing general changes to the way the architecture is operating. The first type of control is faster to activate but lasts only as long as there is a signal present, whereas the second is slower but lasts longer, due to the continued presence of the hormone in the blood. The different types of control require different underlying mechanisms and have different implications for architectural performance.

Within the fields of AI and Cognitive Science the issues of control and representation are somewhat contentious. There is a disagreement between some symbolic and connectionist researchers as to whether control is in fact an issue at all, or just a problem created by using the wrong approach to solving a particular problem (Agre, 1988). Hopefully the above illustration of the types of control present in human beings demonstrates the need to incorporate considerations of control into models of intelligent agents. Perhaps in time connectionist architectures that attempt to model larger and more complex systems will require some effect like hormonal control to allow for more global control changes to be made.

2.4.5 Learning

Learning is vast topic, and arguably central to the adaptation and advancement of life forms, both at a personal and species level (via natural selection). I can't hope to present more than just a very brief look at some of the issues associated with learning that need to be considered in the framework of a resource limitation on architectural development and functionality.

The first point to make about learning is that the scope and variety that an organism is capable of performing depends on its architectural complexity. Learning can vary from simple reinforcement of association through to transformation of representation and control. Various types will provide different benefits, but will also have their architectural costs. Simple reinforcement of association is an acceptable way for an agent to adapt to the demands of the environment, but will never enable it to be able to do something it couldn't do initially. To achieve this, new features of the environment need to be incorporated into existing representations or new representations created (Furse, 1993). Translation of representation allows an agent the ability to integrate new pieces of knowledge, and causal relations about the world into its internal model, and develop abilities based on it.

Learning like every other capability has associated costs, which influence the operation of the agent. One such cost was proposed by Ballard (1992) to be a limit on the number of things that we can do at once. If we do n things at once, then there are 2^n subsets of combinations that have to be considered to facilitate some learning experience about the actions. Hence if too many things are done at once, the number of subsets become too large to operate on. Another cost, or perhaps 'side effect', is the role of learning in psychopathology. Overcoming uncertainty is part of the function of learning. However this requires that the environment a human infant is in, whilst it is developing the representations it needs when adult, needs to be an accurate facsimile of the 'real world'. If the child's world does not have good examples of the skills and relationships it needs to function as an adult, when it grows into an adult it will in some sense be dysfunctional. The dysfunctional behaviours and thought patterns are what underlie some psychopathologies (Beck, 1991).

3 How Systemic Design Overcomes The Problems Of The Design-based Approach

In the introduction to this paper I listed three criticisms that had been made against the design-based approach, as a methodology for studying emotional phenomena (amongst other things). I wish to draw this paper to a close by showing how these criticisms can be dealt with by using Systemic Design. The first point raised concerned the intractability of splitting complex interdependent architectural phenomena down into unitary phenomena to allow them to be modelled and studied. This is an attack on the reductionist stance, and whilst it is true that as Systemic Design incorporates the design-based approach (and hence is also advocating a reductionist stance), its broader mixed top down and bottom up approach offers a partial solution. The scope of the phenomenon(na) under study by the researcher can be specified to correspond to the balance of bottom up or top down design they are using.

If an agent's architectural functionality does exist only in the form of a network of interdependencies which can't be split into unitary phenomena, Systemic Design allows data taken from the real system to be used to set up the skeleton of the network, allowing more detailed modelling of specific parts of it to take place (i.e. a combination of bottom up and top down design, reinforced with data gathered from real organisms). As to the implicit message in the above criticism, that a system like a human being is so complex that the only complete fully functional model of it possible is the system itself, there are insufficient data available at the moment to know for sure. If it is the case then Systemic Design is as futile an approach to understanding it as any that are currently used. However it may provide a methodology for reaching that conclusion faster than other approaches alone!

The second criticism (Clark, 1989) concerned the implausibility of systems designed without the same biological constraints under which we evolved, being able to generate good explanatory theories about human capabilities. Systemic Design deals with that problem by incorporating both phylogenetic considerations and an understanding of the resource limitations that have some influence on a biological life form. It represents a much more tractable approach to producing good theories of mechanism, than going back to attempt to build simplistic neurally plausible models. The problems with that approach stem from the vast number of generations of intermediate life forms between the simple organism and human beings: Firstly there is the combinatorial problem in deciding which characteristics of one successful agent should be carried forward into its successors. Secondly (assuming the first one can be got around), there is the impossibility of being able to travel down the same evolutionary path through evolutionary space that produced a life form like us. Systemic Design allows the researcher to go back just far enough to extract the data needed to under-pin the mechanisms being proposed.

The final criticism (Pfeifer, 1993) suggests that as there is no clear understanding of the relation between mechanism and behaviour, it is impossible to build accurate models of the underlying mechanisms. Systemic Design deals with this point by virtue of its phylogenetic considerations and it holistic design strategy. Phylogenetic considerations can be used to illustrate how evolutionary simpler life forms mapped behaviours to mechanisms, and the holistic design strategy allows for data taken from different research fields (e.g. psychobiology or neuropsychology) to be used to match similar brain structures across evolutionary separated organisms. By considering different mechanisms with different characteristics, different known behaviours, and emergent behaviours can be investigated.

4 Conclusion

In this paper I have shown how the design-based approach can be extended to produce a methodology that does not have its limitations. It is not just a technique that can only be used by researchers trying to generate theories about mechanisms implemented using computer models (although this is its primary purpose). It can also be used by people who would normally use a bottom up design approach to provide a framework within which to refine their data driven models, and test the implications of the theoretical mechanisms they are suggesting.

References

Abrams, P. A. (1986). Adaptive responses of preditors to prey and prey to preditors: The failure of the arms-race analogy. *Evolution*, 40(6):1229–1247.

Agre, P. E. (1988). *The Dynamic Structures of Everyday Life*. PhD thesis, Artificial Intelligence Laboratory, MIT.

Ballard, D. (1992). Communication to aaron sloman.

Beck, A. (1991). *Cognitive Therapy and the Emotional Disorders*. Penguin Books.

Clark, A. (1989). *Microcognition*, chapter 4. MITP.

Clark, D. M. (1986). A cognitive approach to panic. *Behavioural Research and Therapy*, 24(4):461–470.

Dawkins, R. (1986). *The Blind Watchmaker*. Penguin Books.

Frijda, N. H. (1986). *The Emotions*. Cambridge University Press.

Furse, E. (1993). Escaping from the box. In *Prospects for Artificial Intelligence, proceedings of AISB93, the ninth biennial conference of the Society for the Study of Artificial Intelligence and the Simulation of Behaviour*, pages 71–80. IOS Press.

Gazzinga, M. S. (1975). *The Handbook of Psychobiology*, chapter 15,16. Academic Press.

Gray, J. A. (1987). *The Psychology of Fear and Stress*, chapter 5. Cambridge University Press.

Pfeifer, R. (1993). The new age of the fungus eater: Comments on artificial intelligence and emotion. A shortened version of which will be published in the proceedings of ECAL-93.

Read, T. and Sloman, A. (1993). Terminological pitfalls of studging emotion. In preparation.

Sloman, A. (1993a). Personal communication.

Sloman, A. (1993b). Prospects for ai as the general science of intelligence. In *Prospects for Artificial Intelligence, proceedings of AISB93, the ninth biennial conference of the Society for the Study of Artificial Intelligence and the Simulation of Behaviour*, pages 1–10. IOS Press.