Simulating Infant-Carer Relationship Dynamics

Dean Petters,
School of Computer Science,
The University of Birmingham
d.d.petters@cs.bham.ac.uk

Abstract

Advances in autonomous agent technology have resulted in the potential for implementations of multiple agents to act as psychological theories of complex social and affective phenomena. Simulating attachment behaviours in infancy provides a relatively simple starting point for this type of theory development. The presence of neurophysiological, psychological and other types of data facilitates the validation of architectural theories by constraining these architectures at multiple levels. A seven part design process is described which details how requirements are specified and how design, implementation and evaluation processes are carried out. Two competing theories are proposed, one that involves some deliberation and one that is reactive only.

Introduction

This paper describes work in progress that brings together sources of empirical data and methods of validation in an original interdisciplinary manner. The project's aim is to further understanding of infant attachment behaviour using a design perspective. This means explaining what structures would be required in a system's design to enable attachment phenomena to be produced. What is emphasised in this paper are not details of the simulation, but rather the novel relationship between the empirical data, the nature of the simulation and the methods of validation.

The particular attachment behaviour under investigation is the pattern of infant response to separations from and subsequent reunions with their carers in a controlled procedure that occurs in an unfamiliar laboratory environment. This procedure is known as the 'Strange Situation Experiment'. A key finding is the different patterns of infant behaviour found in separation and reunion episodes. An autonomous agent simulation is being produced that will act as an architectural theory of these patterns of behaviour at a high, goal-oriented level. The existence of a wealth of linked empirical data and theory from cross-species, evolutionary, physiological and anthropological branches of Attachment Theory helps constrain and validate the architectural theories formed.

The behaviour to be explained

This section illustrates the empirical observations being investigated.

> "*A mother and her child are in an unfamiliar room together. The child is about a year old and plays with some toys on the floor as the mother reads a magazine. An unfamiliar woman enters and starts to chat to the mother and then tries to play with the child. After a short time, the mother gets up and leaves her child alone with this stranger. What will the child do now? What should the child do now? If the mother returns to the room after a few minutes how will the child respond?*" [After the separation from its mother the child might:] "*cry, throw a tantrum or become anxious, attempt to follow the mother out of the door or carry on as if nothing had happened. In observing infants in naturalistic settings and in this laboratory-based Strange Situation, researchers have witnessed all of these reactions; but taken in isolation, the infant's reactions to such a separation tell us remarkably little. The crucial issue for attachment research is how children react on being reunited with their mothers.*"(page 1 (Meins 1997))
>
> "*Consider children who cry when their mothers leave the room: some of them will be consoled immediately by the mother's presence, others will need to be picked up and hugged before they can be comforted; but some children will cry more angrily when their mothers return and when they are picked up to be comforted may even strike out and stiffen or squirm on their mother's arms. Children who attempt to follow their mothers may also react in a number of ways on reunion: they may show a desire for physical contact and closeness by approaching her or clambering onto her knee; alternatively, they may be content with merely greeting her and continuing to interact with her at a distance. Even children who seem oblivious to their mothers' leaving may surprise us in their reactions on being reunited. Some will carry on playing in much the same way, paying as little attention to the mother's return as they did to her exit, but others will immediately approach and want to be picked up.*" (page 2 (Meins 1997)).

What we want to do is understand the architectural mechanisms by which these behaviours come about, and to form a theory explaining the purpose, if any, of these behaviours for the infant. We are interested in how patterns of behaviour are formed by the interaction of the infant's cognitive architecture and its environment. The separation behaviours are not a clear guide in this regard. This is because separation behaviours in the Strange Situation laboratory setting are not fully predicted by the carer's and infant's behaviour in the home environment. But importantly, reunion behaviour in the Strange Situation is strongly predicted by the home behaviour of the mother and the infant.

Regardless of how they reacted in separation;

- the children whose response to their mothers on reunion was: to not seek contact or avoid their mother's gaze or physical contact with her, *are described as insecure-avoidant and labelled type A*. These children return quickly to play and exploration but do so with less concentration than secure children. Whilst playing they stay close to and keep an eye on their carer. They received care at home which was rated as insensitive, rejecting, and interfering or ignoring.
- the children whose response to their mothers on reunion was: positive, greeting, approaching, making or accepting contact with, or being comforted by her, *are described as securely attached and labelled type B*. These children returned to play and exploration in the room sooner and received care at home which was rated as sensitive, accepting, cooperative and accessible.
- the children whose response to their mothers on reunion was: not being comforted and overly passive or showing anger towards their mothers, *are described as insecure-resistant/ambivalent and labelled type C*. These children do not return quickly to exploration and play. They received care at home which was rated as insensitive, only moderately accepting and moderately cooperative and often ignoring.
- the children whose response to their mothers on reunion was: totally disorganised and confused, *are described as insecure-disorganised and labelled type D*. The home environment of behaviour for this very small proportion of infants with disorganised attachment has been found to be often dysfunctional ((Meins 1997), (Ainsworth *et al.* 1978), and (Weinfield *et al.* 1999)).

The nature of the simulation

The aim is to build infant and carer software agents that reproduce the differing patterns of reunion and separation behaviour found in studies of attachment. These simulated agents should be designed in a manner that increases our understanding of how and why the patterns are formed in reality. The simulation needs to be as simple as possible whilst being powerful enough to represent plausible competing mechanisms and possible causal structures that underlie the patterns of behaviour.

We are looking for the right level of abstraction to represent the problem. For instance the simulation does not replicate the lower level details of sensory modalities and motor actions. What is needed in the simulation are mechanisms of sensation and action that fulfill the appropriate functional role.

We are not interested in simple but implausible solutions such as matching correlated behaviours with look up tables or If...then rules. Any mapping between input data (in the form of home study records) to output data (in the form of the Strange Situation observation records) should emerge from the agent acting towards plausible goals or functions. These might include: managing a balance between safety and learning functions in the home environment or finding the best sub-goals to help gain a sense of security in re-union episodes.

The simulation must be consistent with what is known about infant abilities at age 1 in other psychological domains. For instance our representation should be rich enough to include mechanisms for action selection described in theories of executive function from cognitive psychology. The simulation must also have the power to represent the range of theoretical views regarding the phylogenetic and ontogenetic causal nature of attachment styles that can be found within the attachment literature. There is evidence for a range of causal factors in the development of attachment styles. These include the effect of innate temperament, general warmth of carer response and carer response in situations of perceived threat to the infant. All these factors should be represented as requirements in the simulation.

The entire multi-agent simulation can be thought of as an explanatory theory for the behaviours and causal structures under investigation. The agents in this simulation can be described as broad agents because they possess the complete set of perceptual, action and other subsystems needed to act autonomously within their virtual environment (Bates, Loyall, & Reilly 1991). It is this breadth that will allow incorporation of AI and cognitive psychology techniques and theories in the simulation. They can be described as shallow agents because none of these subsystems is defined in great detail. This shallow aspect allows the simulation to be primarily directed at the level of constructs such as goals and action plans without being submerged in low level detail.

Attachment is a *social* and *affective* phenomenon. Comparison with other *social* or *affective* autonomous agent implementations will help to clarify this simulation's particular nature.

The EOS project simulates Palaeolithic social change (Doran & Palmer 1995). A big similarity with a simulation of attachment is the indirect nature of the source material it is simulating. The archeological data it selects as source material is comparable with the codings of infant behaviour used in a simulation of attachment. Both sources constitute a requirement that any designs proposed must fulfill. In both cases source material needs some interpretation before it can be compared with the results of simulations. In the EOS project simulated prehistoric humans are programmed as rule-based systems. These agents have to gather food and accomplish other tasks either as individuals or in collaboration with other agents. Experiments can vary resource types and locations and showed under what initial condi-

LONGER TERM	INTERMEDIATE	SHORTER TERM
Personality, Temperament, Attitudes, Skills, Emotions such as love, grief, ***Attachment style?***	**Moods, Beliefs, Preferences, Emotions** such as joy, fear, **Intentions, Plans, Desires**	**neural and physical events,**

Figure 1: Classes of semantic control state, compared with respect to the approximate duration that each class of control state may exist as a disposition within an architecture. Adapted from (Sloman 1995).

tions semi-permanent social groups formed. For example, if resources are lowered but also concentrated in a few places and agents possessed sufficient information processing complexity then social hierarchies emerge. Contrasting the EOS project with a simulation of infant attachment we find that the EOS project puts greater emphasis on rational interaction between agents. In attachment more emphasis is put on affective and reflexive responses. In the EOS project time is constrained, but not at the scale of seconds, and when EOS agents negotiate they can't sense detailed aspects of other agents demeanour, such as their direction of gaze. To simulate attachment we need to represent time and modes of communication, such as joint attention, at a finer degree of granularity.

Comparing this simulation of attachment with other affective agent simulations we find much in common. For example the infants in a simulation of attachment may sense and move towards objects of interest much as the agents in Canamero's (Canamero 2002), Allen's, (Allen 2001), Gadhano's (do Carmo Gadhano 1999) or Scheutz and Logan's (Scheutz & Logan 2001) simulations sense and move towards target objects such as food. The temporal granularity of how the agents perceive and act and the representation of affective states such as fear may be similar. What sets it apart from other affective agent simulations is the temporal nature of the attachment states it is attempting to simulate. In line with Attachment Theory ((Bowlby 19691982) chapter 13), this project views attachment styles as longer term affective control states (Sloman 1995) somewhat similar to personality, skills and attitudes (figure 1). Attachment styles can be formed from the experience of low level events and short term control states, but also possession of a certain attachment style predisposes an architecture to particular short term patterns of states such as plans or emotions.

Infants switch between attachment behaviour and other activities, such as exploring their environment. Therefore a simulation of attachment needs to include mechanisms for how infants interact with and learn about objects they perceive. There exist a large number of implemented simulations of cognitive development which do not use autonomous agents but model the workings of a single module that improves its performance on a limited subset of tasks (Klahr 1999) (Elman 1996). A much smaller number of simulations buck this trend in cognitive modeling and represent cognitive development within autonomous agents (Drescher 1991) (McCauley 2002) (Viezzer 2003). The relationship of these simulations of cognitive development with a fully developed simulation of attachment is that a simulation of attachment needs to subsume their mechanisms of concept formation. To represent the development of attachment over time we need to have short term and long term affective states existing alongside structures and processes related to the development of concepts like object and person permanence.

The design process

The seven aspects of the design based approach followed in this work are:

- **Selection of source material**
- **Scenario formation**
- **Design formation**
- **Implementation**
- **Internal evaluation**
- **Evaluation of surrounding design space**
- **External validation**

The nature of these seven parts, and hence the overall process, are evolving as the project progresses. The process has previously been described as a five part process ((Beaudoin 1994) page 5 and (Wright 1997) page 22), but the first and last parts of the list above have been added because of the complexity of interpretating empirical data and validating the whole process. The seven parts can be followed in series in any order, in parallel or some mixture of the two. In general terms, the selection of the source material and the scenario are involved with the setting of a problem. The design and implementation are concerned with the solution to that problem. The internal evaluation and evaluation of the surrounding design space assess how good the solution is to the problem. All these parts are considered by the external validation which is concerned with whether the whole process conforms to rational science.

Selection of source material

The process of interpreting the empirical data includes editing out redundant aspects of the core empirical studies and including relevant details from other fields such as cognitive development and studies of executive function in infancy. Direct observations and general theories can be combined. What is important is that elements of the theory under testing are demarcated from unproblematic background knowledge (Lakatos 1970). Source material is considered unproblematic background knowledge if this view is reached as consensus within the research community. It is within this framework that an analysis of adaptive functions of infant behaviour related to the strange situation provides three goals for behaviours seen in the strange situation. These goals are safety, exploratory learning and social learning, and they are not necessarily explicitly represented by the infant. Rather, evolutionary ethological theory holds that patterns of behaviours to fulfil these goals have evolved in humans and other species (Bowlby 19691982), (Chisholm 1999), (Belsky 1999), (Simpson 1999). A qualification to this conclusion is that the reasoning about the functional nature of the A, B and C attachment styles does not necessarily hold for the D disorganised category of infant attachment behaviour. The very small numbers of infants possessing this attachment type and its pathological nature mean that we can less confidently reason about its adaptive function.

When considering the issue of aetiological theories of infant attachment there is no such consensus within the community of individuals that study causes of individual difference. Therefore temperament based, warmth based and threat based aetiological theories need to be capable of being included. The mix of aetiological factors which contribute to the development of attachment style, such as the effects of learning and temperament, should be represented in some mini-scenarios but may be omitted in others.

Scenario formation

Once the material to be included in the simulation has been decided upon it has to be organised in a manner that allows for precise evaluation. Scenarios are central to the validation of any designs that are implemented. This is because they form a detailed and graduated requirement against which competing implementations can be assessed.

Scenarios should capture abstract patterns of behaviour that are unproblematic with regard to the theory being developed and reform them at a level of concreteness and detail appropriate for the requirements specification of the design and implementation phases. The core of the scenario in this project will be an adaptation of the behaviours described in section 2, but with greater emphasis on a broader range of infant abilities and activities, particularly the infant's experiences in the home environment.

For each A, B and C infant-carer dyad there will be mini-scenarios that detail the two agents behaviour in a prolonged training session (as previously noted the type D dyad will initially be omitted from consideration). This training period is based upon home observation information from attachment studies and will include periods of separation and reunion that vary in number and duration. Then a final 'testing' session will occur with controlled separation and reunion episodes, mirroring the strange situation experiment itself.

The empirical studies of the strange situation show that in reunion episodes nearly all infants do some crying and in separation episodes nearly all infants do some exploration. Behaviours can switch quickly and may oscillate. This state of affairs needs to be reproduced as the required outcome in mini-scenarios. The scenarios will not be in the form of scripts that the agents keep to word for word, like an actor performing Shakespeare. It is the frequencies and durations of actions and intentions over and extended period that have to match.

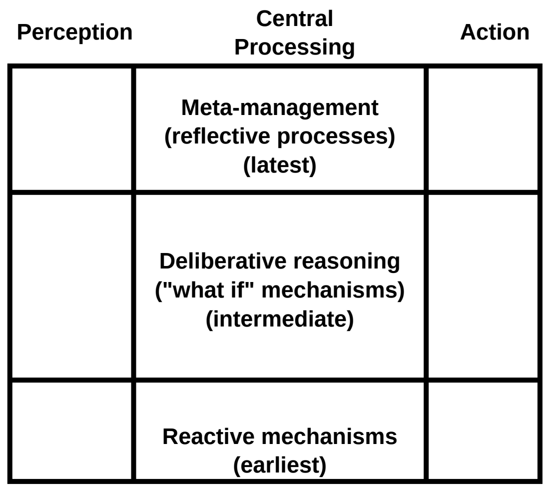

Figure 2: The cogaff schema represents an architecture space, allowing diverse architectures to be systematically compared. In phylogenetic and ontogenetic development, Reactive processes form earliest, and meta processes latest.

Design formation

Since there is not an existing implementation against which a single design can be validated, this design process needs to propose at least two contrasting designs. These should aim to satisfy the scenario in interestingly different ways. Each design needs to describe what component algorithms and representations are required, and how these components are related in an overall architecture.

This section proposes two designs, each of which can adapt under appropriate training schedules to reproduce the A, B and C attachment styles that are observed in the strange situation experiment. Other related designs will be discussed in the evaluation of surrounding design space. The designs are described within a framework provided by the cogaff schema (see figure 2) (Sloman 2001), (Sloman 2002). This is beneficial because the designs are then formulated in a manner that facilitates comparison, and hence validation. The cogaff schema categorises processes according to whether they occur in perception, central processing or action subsystems and whether they occur at reactive, deliberative or meta-management levels. It allows many different patterns of information-flow (including control-flow) between mechanisms located at different parts of the grid.

Features of the design may be inspired by existing theories from the attachment literature, derived from other sources or developed from scratch during the design process. Within the attachment literature there exist some theories which explain attachment phenomena by reference to architectural features of design. These theories are described in linguistic terms only and vary widely in a number of aspects. For example some explain behaviour are derived from animal studies and explain behaviour at a neurobiological level (Kraemer 1992) (Hofer 1995) (Polan & Hofer 1999). Others focus on humans and invoke concepts from cognitive psy-

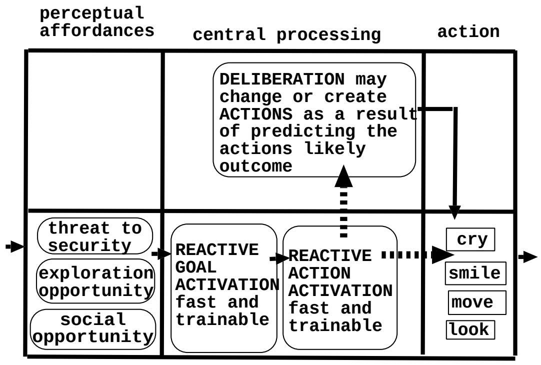

Figure 3: Architecture Design Type 1: A reactive and deliberative architecture to support strange situation patterns of behaviour. The dashed arrows represent routes through the architecture that that compete, and may switch moment to moment during a three minute episode. Reactive training occurs by re-inforcement learning over many prior episodes. Processes related to feedings, obstacle avoidance and exploratory and social learning are omitted from the diagram, and some of these may occur at a deliberative level.

chology (Main, Kaplan, & Cassidy 1985) (Main 1991) (Crittenden 1995) , AI (Bretherton 1990) (Bretherton & Munholland 1999) and control theory (Kobak *et al.* 1993). These theories do not exclude each other, and ultimately all may be reconciled within a broad simulation.

A noteable absence from this set of theories is an explanation for strange situation behaviour early in infancy that relies upon deliberative or meta level mechanisms of executive function that operate before rich linguistic representations develop.

Figure 3 shows a generic architecture of Design Type 1. This design allows reactive and deliberative setting of actions. It produces prepotent actions in response to changes in its environment and also has the ability to deliberate about the consequences of these actions and modify them if needed. Prepotent responses can be learned from previous experience or are innate reflexes that are present in the architecture because they brought about safety over the course of evolutionary development. This design does not possess any mechanisms of meta-management by which its policy of control might be changed and therefore whether deliberation of a given action will occur is decided by a fixed control policy. Meta-management to allow a variable policy in the control of deliberation activation may be added later.

Figure 4 presents an architecture trained to produce secure (B) style attachment behaviour. It shows that if an infant possessing an architecture of type 1 receives sensitive and effective care over an extended period then in reunion episodes it will tend to activate the goal of gaining safety less strongly. Anthropomorphising the infant might be said to *"trust its carer will ensure its safety"*. It will then tend to activate exploratory and social goals more easily and frequently, and hence exploratory and social actions. These actions may be processed via reactive or deliberative routes.

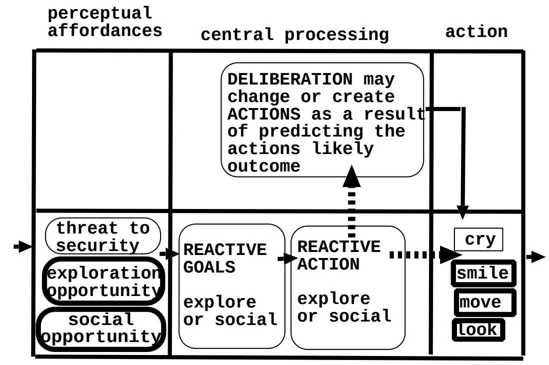

Figure 4: Reunion behaviour in an architecture of Type 1 that has been trained to support type B secure attachment behaviour. Mostly smiling, moving and looking actions occur in reunion episodes.

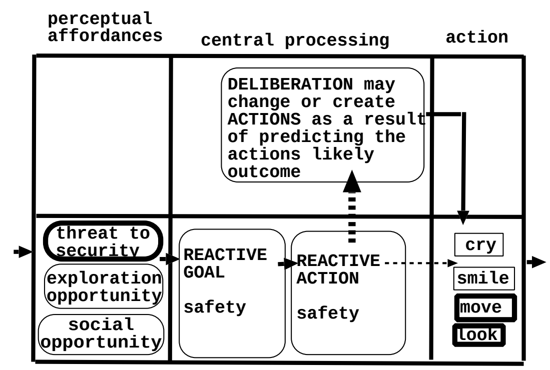

Figure 5: Reunion behaviour in an architecture of Type 1 that has been trained to support type A avoidant attachment behaviour. Mostly moving and looking actions occur in reunion episodes.

An architecture that reproduces avoidant (A) style behaviour is shown in figure 5. If the care an infant receives is insensitive but consistent then in reunion episodes it will tend to activate the goal of increasing safety. From this infants experience the carer is not predicted to act in a gratifying manner when the infant bids for attention. Prepotent 'affective' action will often be deliberated upon and substituted with actions that are predicted to bring about safety without communication of negative affect. Figure 6 shows an architecture to support ambivalent (C) style behaviour. If an infant with the same initial architecture receives insensitive and inconsistent care then in reunion episodes it will tend to activate the goal of increasing safety, and most often follow the prepotent action of communicating negative affect.

An interesting question is whether a totally reactive design can be considered plausible or whether some version of a mixed design will be deemed to be necessary. Figure 7 illustrates an architecture of Design Type 2. This design is somewhat similar to the theory put forward by Crittenden (Crittenden 1995). Where it differs from Crittenden's the-

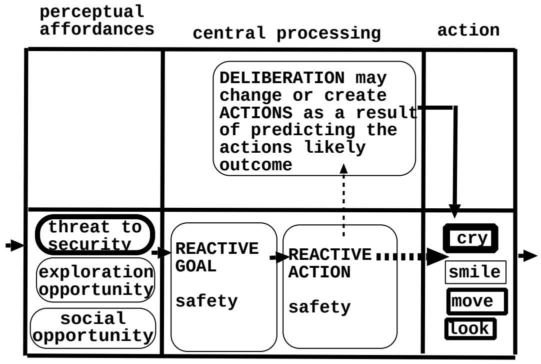

Figure 6: Reunion behaviour in an architecture of Type 1 that has been trained to support type C ambivalent attachment behaviour. Mostly crying with some moving and looking actions occur in reunion episodes.

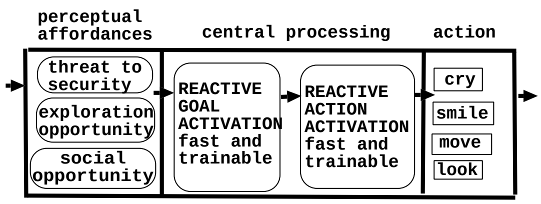

Figure 7: A Reactive Goal Policy Design. The set of reactive design hypotheses postulate that; at one year of age, all Strange Situation patterns of behaviour are produced without resort to deliberative mechanisms. This hypothesis does not preclude simple deliberative processes occurring in other behavioural domains.

ory is that the Type 2 design possesses some extra details. For example the Type 2 design has a two stage nature. First goals are activated and then actions are selected in respect of the active goal.

It should be noted, if the values of variables such as processing speeds, memory capacity, the training schedule and initial knowledge of the agents are not specified, then the performance of reactive and deliberative designs may indistinguishable ((Cooper 2002) page 369). Distinguishing these theories relies upon these variables being set in an implementation.

Implementation

Elman (et al (Elman 1996)) puts the case for implementing designs:

> "*Simulations enforce a rigor in hypotheses which would be difficult to achieve by mere verbal description. Implementing a theory as a computer model requires a level of precision and detail which often reveals logical flaws or inconsistencies in the theory.*"(page 44 (Elman 1996)).

A prototype simulation has been produced in POP11 using the sim-agent toolkit[1]. The simulation environment consists of infant, carer and stranger agents along with food and toy objects. Several infant-carer dyads are situated in separate home spaces and can undergo their training schedules simultaneously. Training and testing sessions can be interacted with whilst they are running, for example by using a mouse to move agents or by making changes to an agent's internal database. This reduces the risk that the implementation works only in special cases. The home areas comprise of food and toy objects with varying positions and attributes which the infants become familiar with by remaining in sensory contact with for a moderate period of time. Carers have the task of collecting energy from food objects and taking this energy to feed to the infants. Carers should also respond with communications, which may vary in affective valence, to social bids from the infants.

The infants possess perceptual systems which assess the environment for whether it affords opportunities for social and exploratory learning and whether it poses a threat to the infants safety. Factors which increase the measure of threat are: the presence of unfamiliar objects (especially unfamiliar agents), objects oriented in unfamiliar ways and a history of numerous previous episodes when the infant's bids for attention from the carer have failed. The affordances for exploratory learning are derived from the distance to objects and the time since the object was last focused on. The affordance for social learning is similarly produced from immediate, opportunistic measures, such as distance to the carer and orientation of the carer, and the cyclical measure of the time since the last social interaction. The affordances compete in a winner takes all manner, the affordance with the highest activation forming a goal that directs the infants actions. Actions in response to threat have an initial innate bias involve negative affective communication, in short, 'crying'. In the implementation of the reactive design this predisposition is extinguished slowly over many episodes by a process of re-inforcement learning. In the implementation of the deliberative design the reactive mechanism may also be present but a deliberative process may also be activated.

At present the simulations have not been constrained by adding bounds on processing speed or memory capacity, and the training schedules and innate knowledge have not been limited or controlled. Addition of these factors is key to the three processes of evaluation.

Internal evaluation

This part of the process is concerned with the internal consistency of each competing design. It iteratively assesses whether each part meets the requirements in relation to other parts of the overall process. The scenario must be a faithful abstraction of the source material, the design must meet the specification provided by the scenario and the implementation must possess a design that matches that already described. After each internal evaluation the implementations and designs are both, if possible, fitted to satisfy their requirements.

[1] http://www.cs.bham.ac.uk/~axs/cogaff/simagent.html

Evaluation of surrounding design space

To fully understand a design we need to understand the space of possible designs (Beaudoin 1994). Of interest are issues such as: how the design might have been different and trade-offs implicated in the design. We want to see the affect of slight changes in the requirements and design. Approaches such as Monte Carlo simulations, computational experiments and comparative modelling may be used to evaluate the design space of simulations (Cooper 2002).

External validation

External validation is concerned with assessing the competing designs and implementations with regard to how good they are as theories of the target phenomena. If possible, we would like to form our theories so that they are falsifiable. According to Lakatos, old (or rival) theories are only falsified when new theories are proposed that explain all the successes of the old, predict additional novel facts and some of these novel facts are empirically corroborated (Lakatos 1970). Following this high standard for falsification means that complex architectural theories may have to have a prolonged period of development before we can falsify them. This is because broad architectures possess many degrees of freedom and ad-hoc and arbitrary corrections can be made at the implementational level. So how do we rationally validate theories if there is no reasonable likelihood of falsifying them in the near future?

Lakatos (Lakatos 1970) provides a solution to this problem by his reconception of static theories into dynamic Research Programmes, which change over time and are evaluated as undergoing progressive or degenerative problem shifts. In the context of this project, progressive problem shifts to simulations may involve adding detail at lower explanatory levels, such as at neurophysiological levels, that are within acceptable principled constraints. Degenerative problem shifts might involve matching the simulation to its required behaviour by ad-hoc, unprincipled hand-coding of details whose only rationale is to reproduce the target behaviour. Attempting to validate architectural theories of information processing in thes terms is not new. For example Newell presents SOAR as a Lakatosian Research Programme (Newell 1990). Lakatos separated the theoretical contents of Research Programmes into two types of information: core assumptions (which identify the theory and are not expected to change) and peripheral assumptions (which form a *"protective belt"* around the core assumptions and change in response to new empirical data). Cooper (Cooper 2002) suggests this relationship is isostructural with the difference between the aspects of architectural design which don't change between differently implemented simulations of that design and implementational details that may vary between implemented simulations. Thus simulations can be validated by assessing the nature of the changes that emerge in the implementational details as the simulation is fitted within empirical constraints.

Conclusion

The design process described above is based upon that previously used within the Cognition and Affect group in Birmingham, for example as described by Beaudoin (Beaudoin 1994) and Wright (Wright 1997). The scenario and design in this work are also somewhat similar to those produced by Beaudoin and Wright but the design process differs from these two previous studies because of the closeness of the mapping it attempts with data drawn from the psychological literature. It accomplishes this because the scenario is created from analysis of a psychological domain. All consequent design decisions are then evaluated with regard to how they improve the simulation as a psychological theory of that domain. From the perspective of Attachment Theory, this work is novel because it goes beyond linguistic descriptions, produces an implemented simulation as a theory of attachment and incorporates mechanistic detail beyond the normal remit of Attachment Theory.

Although the current scope of this work is focused on attachment, the potential scope is broader. This approach will readily allow additional phenomena to be incorporated into the design process. For example the simulation might be extended in the age range it covers, or to cover other domains, such as cognitive tasks, or lower levels of implementation, such as incorporating physiological correlates of attachment into the simulation.

Its high level, top-down, starting point constitutes a meta-theoretical assumption that limits its potential for capturing some types of emergent phenomena, and therefore this approach should be considered the complement of more 'bottom-up' approaches. For example, from the Attachment literature, Main (Main 1999) makes the, as yet unrealised, suggestion to investigate individual differences in attachment using evolutionary computer simulations. From AI, Scheutz and Logan (Scheutz & Logan 2001) present a systematic evaluation of the rival performance of reactive and proto-deliberative mechanisms in a simulated environment that might be adapted to a simulation of attachment.

Acknowledgment

I would like to thank my Phd supervisor, Aaron Sloman, for much help and encouragement.

References

Ainsworth, M.; Blehar, M.; Waters, E.; and Wall, S. 1978. *Patterns of Attachment: a psychological study of the strange situation*. Hillsdale, NJ: Erlbaum.

Allen, S. 2001. *Concern Processing in Autonomous Agents*. Ph.D. Dissertation, School of Computer Science, The University of Birmingham.

Bates, J.; Loyall, A. B.; and Reilly, W. S. 1991. Broad agents. In *AAAI spring symposium on integrated intelligent architectures*. American Association for Artificial Intelligence. (Repr. in SIGART BULLETIN, 2(4), Aug. 1991, pp. 38–40).

Beaudoin, L. 1994. *Goal processing in autonomous agents*. Ph.D. Dissertation, School of Computer Science, The University of Birmingham. (Available at http://www.cs.bham.ac.uk/ research/ cogaff/).

Belsky, J. 1999. Modern evolutionary theory and patterns of attachment. In Handbook of Attachment, *eds. J. Cassidy & P.R. Shaver*. London: Guilford Press. 141–162.

Bowlby, J. 1969/1982. *Attachment and loss: volume 1 attachment*. New York: Basic books.

Bretherton, I., and Munholland, K. 1999. Internal working models revisited. In Handbook of Attachment, *eds. J. Cassidy & P.R. Shaver*. London: Guilford Press. 89–111.

Bretherton, I. 1990. Open communication and internal working: their role in the development of attachment relationships. In Socioemotional development, Nebraska Symposium on Motivation No 36, *eds. R. Thompson*. Lincoln, Nebraska: University of Nebraska. 57–113.

Canamero, D. 2002. Designing emotions for activity selection in autonomous agents. In Trappl, R.; Petta, P.; and Payr, S., eds., *Emotions in Humans and Artifacts*. Cambridge, MA: MIT Press. 115–148.

Chisholm, J. 1999. *Death, hope and sex.* Cambridge: CUP.

Cooper, R. 2002. *Modelling High-Level Cognitive Processes*. New Jersey: Lawrence Erlbaum Associates.

Crittenden, P. 1995. Attachment and psychopathology. In Attachment theory: social, developmental and clinical perspectives, *eds. S. Goldberg R. Muir & J. Kerr*. Hillsdale, NJ: The Analytic Press. 367–406.

do Carmo Gadhano, S. P. C. 1999. *Reinforcement Learning in Autonomous Robots: An Empirical Investigation of the Role of Emotions*. Ph.D. Dissertation, The University of Edinburgh.

Doran, J., and Palmer, M. 1995. The eos project: integrating two models of palaeolithic social change. In Artificial Societies, *eds. N. Gilbert & R. Conte*. London: Guilford Press. 103–125.

Drescher, G. 1991. *Made Up Minds*. MIT Press. (revised edition).

Elman, J. 1996. *Rethinking innateness : a connectionist perspective on development*. London: MIT Press.

Hofer, M. 1995. Hidden regulators in attachment and loss. In Attachment theory: social, developmental and clinical perspectives, *eds. S. Goldberg R. Muir & J. Kerr*. Hillsdale, NJ: The Analytic Press. 203–230.

Klahr, D. 1999. The conceptual habitat: In what kind of system can concepts develop. In Scholnick, E.; Nelson, K.; Gelman, S.; and Miller, P., eds., *Conceptual Development: Piaget's Legacy*. Mahwah, NJ: Lawrence Erlbaum Associates. 131–161.

Kobak, R.; Cole, H.; Ferentz-Gillies, R.; Fleming, W. S.; and Gamble, W. 1993. Attachment and emotion regulation during mother-teen problem solving: A control theory analysis. *Child Development* 64:231–245.

Kraemer, G. 1992. A psychobiological theory of attachment. *Behavioural and Brain Sciences* 15:493–541.

Lakatos, I. 1970. *Criticism and the Growth of Knowledge*. New York: Cambridge University Press.

Main, M.; Kaplan, N.; and Cassidy, J. 1985. Security in infancy, childhood and adulthood: A move to the level if representation. In Growing points of attachment theory and research, *eds. I. Bretherton and E. Waters*. London: Monographs of the society for research in child development, 50(1-2 serial no 209). 66–107.

Main, M. 1991. Metacognitive knowledge, metacognitive monitoring, and singular (coherent) vs multiple (incoherent) model of attachment: findings and directions. In Attachment across the lifecycle, *eds. C.M. Parkes & J. Stevenson-Hinde & P. Marris*. London: Routledge. 127–159.

Main, M. 1999. Epilogue. attachment theory: Eighteen points. In Handbook of Attachment, *eds. J. Cassidy & P.R. Shaver*. London: Guilford Press. 845–887.

McCauley, T. 2002. *Neural schemas: Towards a comprehensive mechanism if mind*. Ph.D. Dissertation, School of Computer Science, The University of Memphis.

Meins, E. 1997. *Security of attachment and the social development of cognition*. Hove: Psychology Press.

Newell, A. 1990. *Unified Theories of Cognition*. Harvard University Press.

Polan, H. J., and Hofer, M. 1999. Psychobiological origins of infant attachment and separation responses. In Handbook of Attachment, *eds. J. Cassidy & P.R. Shaver*. London: Guilford Press. 162–180.

Scheutz, M., and Logan, B. 2001. Affective vs. deliberative agent control. In et al., C. J., ed., *Proceedings Symposium on Emotion, cognition and affective computing AISB01 Convention*.

Simpson, J. 1999. Attachment theory in modern evolutionary perspective. In Handbook of Attachment, *eds. J. Cassidy & P.R. Shaver*. London: Guilford Press. 115–141.

Sloman, A. 1995. What sort of control system is able to have a personality? (Presented at Workshop on Designing personalities for synthetic actors, Vienna, June 1995).

Sloman, A. 2001. Beyond Shallow Models of Emotion. *Cognitive Processing: International Quarterly of Cognitive Science* 2(1):177–198.

Sloman, A. 2002. How many separately evolved emotional beasties live within us? In Trappl, R.; Petta, P.; and Payr, S., eds., *Emotions in Humans and Artifacts*. Cambridge, MA: MIT Press. 29–96.

Viezzer, M. 2003. The autonomous acquisition of affordance concepts: Exploring a neighbourhood in design space. Technical Report CSRP-03-3, School of Computer Science, The University of Birmingham.

Weinfield, N.; Sroufe, L.; Egeland, B.; and Carlson, E. 1999. The nature of individual differences in infant-caregiver attachment. In Handbook of Attachment, *eds. J. Cassidy & P.R. Shaver*. London: Guilford Press. 68–88.

Wright, I. 1997. *Emotional agents*. Ph.D. Dissertation, School of Computer Science, The University of Birmingham. (Available online at http://www.cs.bham.ac.uk/research/cogaff/).