Implementing Free Will

Bruce Edmonds
Centre for Policy Modelling
Manchester Metropolitan University
http://cfpm.org/~bruce

"*Anyone who considers arithmetic methods of producing random digits is, of course, in a state of sin*" John von Neuman[1]

"*The demonstration that no possible combination of known substances, known forms of machinery and known forms of force, can be united in a practical machine by which man shall fly long distances through the air, seems to the writer as complete as it is possible for the demonstration of any physical fact to be.*" Simon Newcomb, Professor of Mathematics, John Hopkins University, 1901

Abstract

Free will is described in terms of the useful properties that it could confer, explaining why it might have been selected for over the course of evolution. These are: exterior unpredictability; interior rationality; and social accountability. A process is described that might bring it about when deployed in a suitable social context. It is suggested that this process could itself be of an evolutionary nature – that free will might "evolve" in the brain during development. This mental evolution effectively separates the internal and external contexts, whilst retaining the coherency between individual's public accounts of their actions. This is supported by the properties of evolutionary algorithms and possesses the three desired properties. Some objections to the possibility of free will are dealt with by pointing out the *prima facie* evidence and showing how an assumption that everything must be either deterministic or random can result from an unsupported assumption of universalism.

KEYWORDS

Free will, evolution, context, randomness, predictability, rationality

INTRODUCTION

In this chapter I wish to show how free will might be brought about and how this fits into the developmental, social and evolutionary context of the entities that most clearly exhibit this ability: adult humans. That is, I start from the ability that developed humans seem to have and consider how it might be brought about.

In a way I would like to say that there is a *mechanism* for free will, but our archetypal pictures of mechanisms and free will are so inimical to each other that juxtaposing them almost forces a choice between them. That is to say, it *seems* we must choose between either (A) that there is free will but this is not implemented in any mechanism, for it is the nature of mechanisms that they are predictable or (B) that there is no free will because whatever there is must be implemented in a mechanism

[1] Reportedly said by von Neuman at a conference on Monte Carlo methods in 1951

of some kind. I will argue that we *can* have both free will and a mechanism to implement it. I perform this trick by arguing that our picture of mechanisms is inadequate and that we make the mistake of confusing what we can model or understand and the possibilities inherent in what we are trying to model. In particular I will argue that whilst modelling the world using either deterministic or (effectively) random processes is common, that this does not mean that the world is so composed.

For most of this chapter I will assume that free will is possible and thus that it is somehow brought about. The question that I will focus on is *how* it might be brought about. It is, of course, always possible for someone to simply *assert* that free will is impossible. It is my experience in discussing implementing free will that there are some people that just can't abide such a project because it is part of their worldview that free will *must be* impossible – if you are one of these people I suggest that you read the last section, which deals with such philosophical arguments, before reading the rest. If you are prepared to conceive that free will might exist, then I suggest that you read the chapter in the order it is presented (of course the later kind of person has a *choice*).

MODELLING AND CONTEXT

All usable modelling is context-dependent. That is to say that a model (or theory) will have a scope which is the set of circumstances under which the model works[2] and that this scope is related to the context within which the model was developed and validated (though not necessarily restricted to this). Even in physics there is (as yet) no *universal* theory: quantum physics holds for microscopic events, Newtonian physics for events at an everyday scale etc.[3] Even where a theory is said to hold "in theory" the conditions under which it is *usefully* applicable do not always pertain – thus if you are in the natural world with no instruments or tools then Einstein's theory of relativity will not help you understand or predict your environment. For more mundane models such as the ones that capture the movement of a unforced, frictionless pendulum or the process of protein manufacture in the cell, the relevant context is far more obvious (though not necessarily made explicit).

When something interferes from outside the modelling context we often use a 'proxy' for this in the form of an effectively random input (e.g. a "pseudo-random" generator). This is often the best we can do since we cannot extend the model to capture what is beyond the modelling context, but a random source at least mimics the extra-contextuality of this interference. Thus *from within the context* causal factors are often either encodable as explicit parts of a model or represented as random. However this does not *make* this interfering cause random in any other sense (i.e. in different or wider context). The functional properties of free-will mean that such a modelling approach might well result in being represented, from the point of view of an exterior context, by a combination of determinism or randomness. However this does not mean that this is the nature of any mechanism within the brain, i.e. in the interior context. Thus I next look at some of the functional properties of free will.

---

[2] Of course, what 'works' depends on how useful the model is in furthering one's goals for it, so the nature of a model's scope is more complicated than I indicate here.

[3] this is somewhat of a simplification for quantum effects can have consequences in the macroscopic world etc., however it is true that each theory does have its own, non-universal domain of applicability.

A FUNCTIONAL DESCRIPTION OF FREE-WILL

Free will occurs to different extents in different circumstances. It has, presumably, developed in our species as that species has evolved (free-will is something that distinguishes us from, for example, uni-cellular organisms). Thus it is likely that free will has given us some selective advantage otherwise it is hard to see why it would have arisen. I suggest that the properties of free will that are relevant because they have the potential to provide such selective advantage include:

1. *Exterior Unpredictability*. From the point of view of a competitor the actions of the individual possessing free will are (at least somewhat) unpredictable.
2. *Interior Rationality*. One's actions lead to one's goals. That is, from an internal view the actions are consistent with trying to achieve the goals and tend to work towards achieving these goals.
3. *Social Accountability*. When requested, the individual can produce an explanation for the (previously somewhat unpredictable) action in terms that establish its rationality to others. That is a public account of ones decision process can be made so that it can be seen how ones actions were related to one's goals.

The advantages of these arise (on the whole) in social contexts, but it is in social contexts that humans predominantly exist. These social contexts are unavoidable because the survival of humans seems to come from their ability to inhabit many different ecological niches due to their social adaptivity (Reader 1990).

In a partially competitive social situation, where it would be to your competitors' advantage to "guess" what you will do, there is obvious advantage in not being completely predictable. At the same time one needs to perform actions that will further one's own goals. Properties (1) and (2) are trivially easy to reconcile in the special case where the 'rational' thing to do is something effectively random. For example, the direction an animal may bolt when startled. However this is not the case with most human actions which need to be (and are) far more structured.

Membership of many human social groups and institutions (in the widest sense) is often conditional on being able to demonstrate that one is rational (from the viewpoint of the others in that group) so that they can have some assurance that you will abide by the norms and rules of the group. The reason for this is that incentives and sanctions on the way you behave will probably have some sway over you. You do not let a mad person into your home, not because they are more likely to be more violent than a sane person but because any of the usual norms and sanctions one might use to constrain behaviour may have no effect. For example, shaming such a person out of acting dangerously may not work.

Thus we see that simultaneously possessing abilities (1), (2) and (3) is potentially advantageous for us humans (and to a lesser degree other highly social animals). If you lacked (1) you might be predicted and hence out-competed; if you lacked (2) you would be unlikely to achieve any of your goals; and if you lacked (3) you would probably be excluded from many social situations which would otherwise benefit you. Of course, this is a little circular, for many human social structures depend upon the fact that we have properties (3) and (2) and would not be needed at all if it were not for (1). This is unsurprising as this may well have occurred as a result of the co-evolution of our social structures and abilities, as has been suggested for language (Deacon 1992).

Criteria (1), (2) and (3) thus form our requirements for an implementation of free will. It is notable that these criteria are each about different contexts. (1) concerns only the external viewpoint of a competitor; (2) is only about the internal, cognitive context; and (3) concerns the translation of the internal into the external context by an individual.

A PROPOSED MECHANISM FOR FREE-WILL

The key question is *how can these requirements be reconciled via mechanisms that might occur in a brain*. We get a clue from human ontogeny. Free will emerges during development (an human adult has free-will whilst a single-celled foetus does not in any meaningful sense). This is not an instantaneous, all-or-nothing ability but one that *develops* in us over time. As we develop it becomes helpful to have created an internal context that is 'insulated' from outside inspection so that other's can not predict what one is going to do, but sufficiently coherent with others so that social presentations of the content of the internal contexts are judged as reasonable by others.

I suggest that the brain has evolved over millions of years so as to facilitate the development of free will as it develops into adulthood. Thus the proposal is to implement free will using an suitable developmental process. The type of process I have chosen is itself of an evolutionary nature – this corresponds to the suggestion that biological evolution has resulted in the ability to maintain an evolutionary process in the brain – a process that results in the mature brain being able to direct behaviour to the advantage of the individual. The suitability of an evolutionary learning process as the engine of free-will is suggested by the properties of such processes as established in the field of Evolutionary Computation, in particular of Genetic Programming (GP). *Firstly*, GP is a *creative* technique, often coming up with unexpected solutions (Koza et al. 1999). *Secondly*, the mechanism of sexual recombination in GP acts to maintain the *maximum variety* in the population as it learns, since it preserves the subtrees but tries different combinations (Koza 1992, 1994). *Thirdly*, within the space of solutions there are a large set of different solutions that will result in the same behaviour within a particular training context, but will diverge arbitrarily in other contexts (Gathercole 1998). *Fourthly*, it does come up with solutions that match the goal, which is implicit in the selection mechanism (Koza 1992, 1994). *Lastly* there is some evidence that it does facilitate the evolution of evolvability (Altenberg 1994).

Thus I suggest adapting mechanisms from GP to produce a model of the mind that meets the three criteria (1), (2) and (3). The proposed mechanism is as follows:

?? there is a 'space' of possible strategies that are constructible from a 'language' of steps, conditionals and actions – this could be formal language as in GP but does not need to be;

?? there is a current 'population' of strategies from that space that are being evolved as the result of experimental variation of these and their evaluation (based upon the success of using the strategies or similar strategies) – on e can think of these as representing the range of alternatives that one considers in making a decision;

?? the language of these and many of the original archetypes for these strategies have a social origin, i.e. that primitives, the modes of their combination and their meanings are socially shared;

?? the language must be suitably 'open-ended' – that is strategies that are similar in terms of effect must be expressible in many different ways and there must not be a hard upper bound on their complexity;

?? that the success of strategies will be according to their effectiveness in social situations.

This basic structure is augmented in two ways. *Firstly*, by adding the ability to anticipate the results of as strategies and thus allow the evaluation of strategies by whether they produced the expected results as well as the extent to which they furthered goals and, *secondly*, by allowing the (limited) co-evolution of the evolutionary operators themselves. Thus there are the following additional components:

?? the strategies are associated with anticipations of their effects, so that they can be evaluated in terms of whether they produced the anticipated effect as well as their effectiveness in furthering the goals of the agent. This is illustrated in Figure 1 below. For more on algorithms that implement and use such anticipation see (Butz et al 2003).

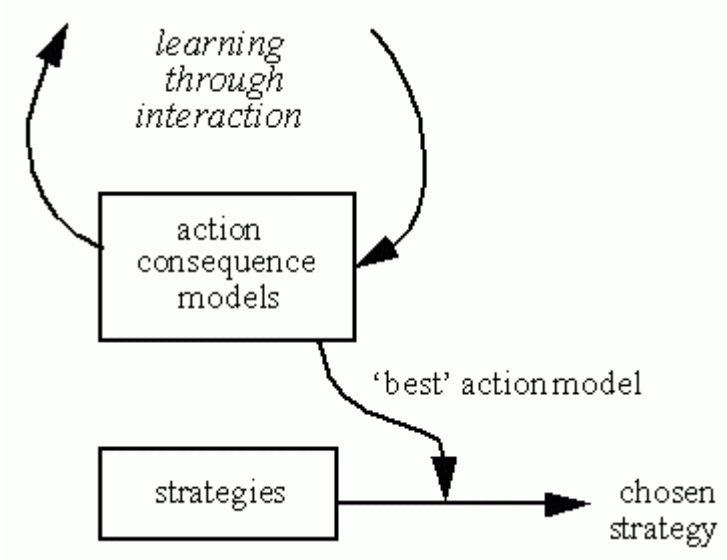

Figure 1. Adding anticipation to the development of strategies

?? the operators that act upon the population of strategies and their anticipations to produce new variations are themselves evolved – the operators act upon themselves to produce variations of themselves and they are selected (at least partially) according to the extent to which they maintain the unpredictability of these strategies. This is illustrated in Figure 2 below. An example of these sorts of algorithm and their properties are discussed in (Spector 1996, Edmonds 2001).

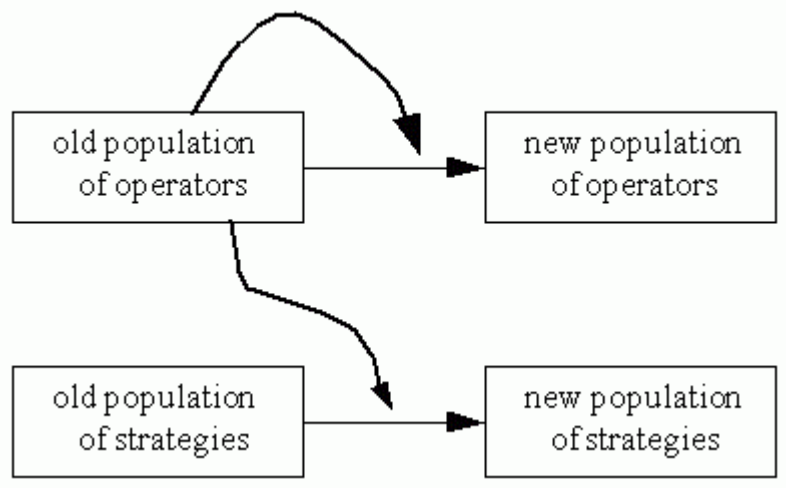

Figure 2. Co-evolving the operators of variation

Obviously in the human case this is, at best, a mere simulaculum of the complexity of human thought and development. For example, it is becoming increasingly clear that the human brain is fundamentally context-dependent in its processing (refs). It would be possible to enrich the above outline to include all of these (for example by including elements such as those described in (Edmonds 2001). However this is beside the point of this chapter – I do not claim that the detail of the above is what happens in the human case, although I do suggest that some evolutionary process like it might be occurring.

These suggested mechanisms, when used in a socially embedded individual will satisfy my criteria. I review each in turn.

1. *Exterior Unpredictability*.

From the point of view of someone else what is observable is the behaviour and reports of other actors. From these observations one may attempt to infer (or, more accurately, guess) the strategy that the individual was attempting to pursue and from that predict the future behaviour. However if there are many different strategies that will produce the same behaviour as that observed then, when the situation changes, the difference between the guessed underlying strategy and the actual strategy may result in different behaviours being exhibited than those predicted from the inferred strategy. The sharpness of the drop in predictive accuracy with a change in context depends upon the algorithms which are used for the development of the internal strategies and for inferring other's strategies. This context-dependency is especially sharp in Genetic Programming. In (Jannik 1994) it is shown how GP can be used to produce random sequences by co-evolving two populations of programs: the first population selected dependent on success at being predicted by the best of the second population and the second selected on the basis of success at predicting the first. The result is a modelling "arms race" with the two populations constantly evolving.

2. *Interior Rationality*.

That an individual should act so as to further its own goals is unsurprising. However it is more difficult to see how this can be maintained in the presence of the drive to (exterior) unpredictability implicit in the account above. The answer comes from the GP structure – there are many different strategies which will result in the same behaviour in any defined set of circumstances, this comes from the redundancy and open-endedness of the language of strategy expression. All of these strategies can be

rational in the sense of furthering the individual's goals. However in new circumstance (i.e. those not in the defined set), these different variations might well result in very different behaviour. Further new developed variations on these different strategies might also result in very different behaviour, even back within the established set of circumstances. This 'sharp context-dependency' in strategy is a consequence of the GP learning algorithm (Altenberg 1995).

3. *Social Accountability*.

The ability to produce an account of why one took any particular action that establishes to the satisfaction of one's peers that one's actions was rational comes out of the shared language of strategies and the underlying rationality of the chosen strategy. Clearly in the human case the formation of the self includes significant social elements. Elsewhere (Edmonds, in press) I argue, that we use our models of others that we infer from our observation of them as a basis for our self-models, and *vice versa*. If this is the case this would provide a deeply shared basis for action strategies. However. this is not necessary for these proposals, for here it is sufficient that the internal strategies can be expressed in a shared language when restricted to the relevant social context.

This proposal also answers the objection that the existence of free will presupposes an infinite regression back in time. That is, free-will is only possible if the decision mechanism and the previous state is freely chosen, which is in turn only possible if the mechanism and the previous state before that is freely chosen etc. However the suggestion that free will evolves in the brain as a human grows up meets this objection –free will emerges by a sort of bootstrapping process in a way that is analogous to how life developed. In such an evolutionary process if you try and chase any particular decision backwards in time then you merely increase the difficulty of modelling it, so that this becomes impractical. The roots of decisions are lost back in the evolutionary process – in a sense, this process can be seen as an way of amplifying the difficulty of modelling (from an external point of view) from small difficulties to insurmountable ones. A rather simplistic way of saying this is that infinitesimal amounts of free will are amplified up to effective amounts by the evolutionary process (just as tenuous and primitive forms of life that are barely distinguishable from mere chemistry has developed into the plethora of life forms we know of today). The strangeness of this "amplification from the infinitesimal" way of expressing the process results from the attempt to impose a context-independent account on a process which effectively creates a new context.

THE SOCIAL AND COGNITIVE VIEWS

This picture of the relationship between the human mind and its social context as a partial explanation for human intelligence is the core of the "social intelligence" (Kummer et al. 1997) and "Machiavellian Intelligence" (Byrne and White 1988) hypotheses. In particular, the latter version is almost an inevitable consequence of the processes I have suggested above. The dual needs of making one's actions unpredictable and continuing to further one's own ends, effectively involves masking one's true intentions – a Machiavellian without such qualities would not be very successful! Yet criterion (3) softens this, since it requires that any such action must be justifiable in a socially acceptable way. Thus the above account of free will is that it would occur and have meaning in a social context.

That the brain implements such a process (or something equivalent) is not clear. That it *could* implement such a process is indicated by the discovery that it does utilise evolutionary processes as part of its functioning (Edelman 1992).

SOME OTHER VIEWS

Most other views on this that consider the possibility of free will are *compatibilist*. That is to say they see free will as compatible with determinism. Many of these seem to take this position due to their commitment to determinism and their experience of free will. My view is that determinism is simply false for it presupposes that models can always be expanded to include extra-contextual interference as fixed rules inside the model.

Aaron Sloman (Sloman 1993) and McCarthy (2002) both argue that it is sufficient that an agent is able to consider and decide what it will do without undue exterior constraint. This is part of the picture but would be inadequate if the behaviour of that agent was predictable to others. If you could guess or infer how that agent reasoned and what factors it based it's reasoning on then, from your point of view (a view exterior to the agent), you could consider that agent as a deterministic machine.

Dennett (1984) also argues in this way, but points out the basic irrelevance of "strong" versions of free will that philosophers object to, arguing that all important aspects of free will, from the point of view of the actor concerned, are possible. Thus Dennett manages to sidestep the issue of context-dependency, allowing the philosophy to cling on to its supposed universality but at the cost of irrelevance to any of the *real*[4] issues concerning human choice. Thus he misses the role that the mental evolutionary process has in separating internal and external contexts and how this essential context-dependency lies at the root of the phenomenon of free will.

PHILOSOPHICAL OBJECTIONS TO THE EXISTANCE OF FREE WILL

Such a proposal as this inevitably provokes many philosophical responses. They are mostly variants of an *a priori* conviction that free will is impossible, and so my suggestion "must be" inadequate.

The most easily dismissed of these comes out of an assumption that the world and hence human decision making is deterministic, despite the fact that evidence does not support this. The spectacular success of quantum physics tells us that, at least at the atomic and sub-atomic level, matter is fundamentally indeterministic. A variation of this is that macroscopic events are, in practice, deterministic – that is to say that for large ensembles of atoms (e.g. humans) the more similar the past situation the more similar will be the outcome. This is not a practical argument since, *firstly*, the degree to which the circumstances must be identical in order that the same human will always decided the same action is far beyond what can ever be arranged and *secondly* these circumstances must include the whole memory of the individual, and hence all the individual's past circumstances. Thus this convergence of human choice with convergence in circumstances is a theoretical assumption (via analogy with simple systems like billiard balls) rather than a matter of *evidence*. The extent of the

[4] Dennett, being a good philosopher, does not of course use the word "real" in this sense but rather such as "practical", "relevant" and "important"

evidence and the possible outcomes in the situation of complete convergence is illustrated in Figure 3 below.

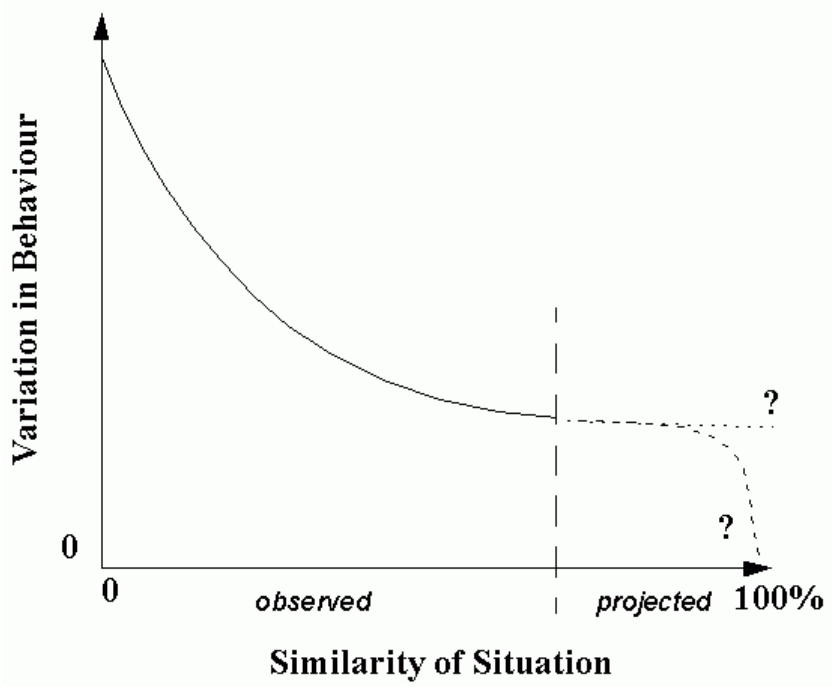

Figure 3. An illustration of the supposed convergence of behaviour with the similarity of situation

A more sophisticated response is that the world is either deterministic or random, and so human decisions must be the same. However, as I argue at the beginning, this is a result of the context-dependency of our modelling, we use randomness as a model of what we can not model and we impute this upon the parts of human decision making we can not model. That we cannot model it is unsurprising since this is (part of) the purpose of free will – to separate the modelling from an internal and external point of view. Few experts on human behaviour will claim that human behaviour is either rigidly determined or random *in practice*, so the insistence to the contrary is more in the nature of a theoretical commitment. The fact that humans (on the whole) can produce a credible reason for their actions makes attributing randomness as a significant cause difficult to maintain and yet there are many cases where humans retain the facility to surprise.

Some (e.g. many economists) will assert that however the individual *does* decide on action, that taken *en masse* their actions are *in effect* random. Even if this *were* the case[5] this would not mean that an individual's action *was* random, merely arbitrary and uncoordinated with others' actions.

There is the objection to free will mentioned above, that free will presupposes prior free will. That is to say for free will to be possible *now* it must have been possible *previously* as the present choice is based upon past memories and the decision mechanism so if these are not free nor can the present choice. The conclusion from this observation is often that free will is impossible. However the proposed

[5] which it isn't, for a consequence of this would be that in larger and larger social systems the randomness would tend to cancel out as a proportion of the total system by the law of large numbers and the system become more predictable, but this is not supported by observation of such systems

mechanisms directly deals with this. This argument does not rule out the possibility of free will, any more than a similar argument would rule out the possibility of life itself. Rather it shows that there is some recursion here, just as life is involved in determining and producing new life, so it is possible that free will is involved in determining and producing more free will. Just because life presupposes prior life to beget it in the process of reproduction, does not mean that life does not exist. It was this thought that first suggested an evolutionary process to me.

Whilst not being simply a matter of degree, there is *ultimately* no hard and fast boundary between having free will and not having it. Like intelligence it is a complex ability, which may be compared only via a sort of sloppy 'short-hand' (A is more intelligent than B). Most people do not have any problem with the idea that intelligence develops with the individual. This bothers some thinkers who wish to *insist* that it is an "all-or-nothing" property, without evidence that this is so but more, I would guess, on the grounds that it enables them to create arguments such as the recursive one above. Hofstadter puts it nicely when he says:

> "*Perhaps the problem is the seeming need that people have of making black-and-white cut-offs when it comes to certain mysterious phenomena, such as life and consciousness. People seem to want there to be an absolute threshold between the living and the nonliving, and between the thinking and the 'merely mechanical,'…*" (Hofstadter 1985)

This brings us to the most fundamental difficulty: it is part of the nature of philosophy to seek universal (non-context-dependent) models for the world. This universalism is partly due to the widespread and accepted practice of using counter-examples in philosophical argument, however weird and extreme these counter-examples are, for this leaves any context-dependent proposal as vulnerable. This tendency leads to an, in practice, assumption that such a universal view is possible once the details of particular contexts are 'filtered out'. If the mechanisms of free will are based in their effectiveness at separating internal and external contexts (from the point of view of modelling) then its existence and the applicability of a philosophical approach are opposed (to the extent philosophy is limited to potentially non-context-dependent arguments or truths). One is then left with a choice: accept that such free-will can exist, for which there is some evidence (albeit mostly anecdotal), or rely on the universality of philosophy (which is pure assumption since there can not be evidence for this).

ACKNOWLEDGEMENTS

Thanks to the participants of the AISB symposium on "How to Design a Functioning Mind" in Birmingham in April 2000 for their discussion and comments upon the first version of this paper (Edmonds 2000a).

REFERENCES

Altenberg, L. (1994). The Evolution of Evolvability in Genetic Programming. In Kenneth E.K. (ed.) Advances in Genetic Programming. Cambridge, MA: MIT Press, 47-74.

Altenberg, L. (1995). Genome growth and the evolution of the genotype-phenotype map. In Banzhaf, W. & Eeckman, F.H. (eds.) Evolution as a Computational Process. Berlin: Springer-verlag, 205-259.

Butz, V.B., Sigaud, O. & Gérard, P. (eds.) (2003). Anticipatory Behaviour in Adaptive Learning Systems. Springer, Lecture Notes in Artificial Intelligence, 2684.

Byrne, R. W. & Whiten, A. (eds.) (1988). Machiavellian Intelligence: social expertise and the evolution of intellect in monkeys, apes, and humans, Oxford: Clarendon Press.

Deacon, T. (1992) Brain-Language Coevolution. Reading, MA: Addison-Wesley.

Dennett, D. C. (1984). Elbow Room: varieties of free will worth having. Oxford: OUP.

Edelman, G. M. (1992). Bright air, bright fire: on the matter of mind. London : Penguin.

Edmonds, B. (1999a). Capturing Social Embeddedness: a constructivist approach. Artificial Behavior, 7(3/4), 323-348.

Edmonds, B. (1999b). Gossip, Sexual Recombination and the El Farol bar: modelling the emergence of heterogeneity. Journal of Artificial Societies and Social Simulation, 2(3), <http://www.soc.surrey.ac.uk/JASSS/2/3/2.html>.

Edmonds, B. (2000a). Towards Implementing Free Will. , AISB'2000 symposium on "How to Design a Functioning Mind", Birmingham, April 2000. http://cfpm.org/cpmrep57.html

Edmonds, B. (2001b) Learning Appropriate Contexts. In: Akman, V. & al. (eds.) Modelling and Using Context - CONTEXT 2001, Dundee, July, 2001. Lecture Notes in Artificial Intelligence, 2116:143-155.

Edmonds, B. (2001c). Meta-Genetic Programming: Co-evolving the Operators of Variation. ELECTRIK, 9:13-29.

Edmonds, B. (in press) The Social Embedding of Intelligence - Towards producing a machine that could pass the Turing Test. In Peters, G. & Epstein, R (eds.) The Turing Test Sourcebook: Philosophical and Methodological Issues in the Quest for the Thinking Computer. Kluwer.

Gathercole, C. (1998) An Investigation of Supervised Learning in Genetic Programming. PhD Thesis, University of Edinburgh, 207 pages. ftp://ftp.dai.ed.ac.uk/pub/daidb/papers/pt9810.ps.gz

Hofstadter, D. R. (1985). Analogies and Roles in Human and Machine Thinking, In Metamagical Themas, New York: Basic Books.

Jannink, J. (1994). Cracking and Co-evolving randomList. In Kinnear, K.E. (ed.) Advances in Genetic Programming, Cambridge, MA: MIT Press, 425-444.

Koza, J.R. Genetic Programming II: Automatic Discovery of Reusable Programs., Cambridge, MA: MIT Press, 1994.

Koza, J.R. Genetic Programming: On The Programming of Computers by Means of Natural Selection, Cambridge, MA: MIT Press, 1992.

Koza, J.R., Bennett , F.H., Andre, D. & Keane, M.A. (1999). Genetic Programming: Biologically Inspired Computation that Creatively Solves Non-Trivial Problems. In Landweber, L., Winfree, E. & Lipton, R. (ed.) Proceedings of DIMACS Workshop on Evolution as Computation, Springer-Verlag, 15-44.

Kummer, H., Daston, L., Gigerenzer, G. & Silk, J. (1997). The social intelligence hypothesis. In Weingart et. al (eds.), Human by Nature: between biology and the social sciences. Hillsdale, NJ: Lawrence Erlbaum Associates, 157-179.

McCarthy (2002) Deterministic Free will. http://www-formal.stanford.edu/jmc/freewill2.htm

Reader, J. (1988) Man on Earth. London: Collins.

Sloman, A. (1992). How to Dispose of the Free-Will Issue. AISB Quarterly, 82:31-32.

Spector, L. (1996). Simultaneous Evolution of Programs and their Control Structures. In Angeline, P. & Kinnear, K. (Eds.), Advances inGenetic Programming 2. Cambridge, MA: MIT Press.