

Unorthodox Optimal Foraging Theory

Anil K Seth

Centre for Computational Neuroscience and Robotics
and School of Cognitive and Computing Sciences

University of Sussex
Brighton, BN1 9QG, UK
anils@cogs.susx.ac.uk

Abstract

As the Simulation of Adaptive Behaviour (SAB) field continues to mature, it is essential that general methodological positions become elaborated into practical programmes of research. This paper describes how a particular flavour of SAB modelling - the use of genetic algorithms to design situated agent (animat) architectures - can effectively complement ‘optimal foraging theory’, as it is understood in theoretical biology. This allows several fundamental problems that arise directly out of the framework of orthodox optimal foraging theory to be addressed, but, as with any trade-off, is not without disadvantages of its own.

1. Introduction

This paper considers the methodological potential of a set of models at one intersection of the overlapping fields of artificial life (AL) and the simulation of adaptive behaviour (SAB). These models, which for want of an appropriate existing term shall be called Individual-based Optimal Situated (IOS) models, derive from the larger class of ‘individual-based models’ (IBMs) and distinguish themselves in two ways. First, they are endowed with a dimension of optimality through the use of genetic algorithms to design agent architectures. Second, they deal exclusively with *situated* agents. ‘Situation’ has a particular meaning here; it connotes not just that agents are contextualised in some way (as will be the case for many IBMs), but that they interact with a spatiotemporally structured environment by means of perception and action. A point of agreement between IBMs and IOS models is that both can be construed as alternatives to ‘classical’ models that consist primarily of state variables and systems of equations that average over properties of individuals.

The task that this paper sets out to achieve is the elaboration of a productive methodological relationship between IOS models and classical ‘optimal foraging theory’ (OFT) models in biology in which IOS models are *not* touted as ‘better’ versions of, or exclusive alternatives to orthodox OFT models, rather, as effective complements.

The essence of the argument is as follows. Orthodox OFT (introduced in Section 2) is an attempt to understand both functional and mechanistic aspects of animal (foraging) behaviour by treating observed behaviour as ‘optimally’ adapted to an environmental ‘problem’. I argue that orthodox (classical, equational) OFT models carry with them several substantial ‘framework assumptions’ which both limit their explanatory flexibility, and undermine the rationale behind the ‘incremental revision’ process that characterises the orthodox OFT research cycle. I concentrate (in Sections 3 and 4) on identifying the problematic consequences of these framework assumptions and indicating in each case how IOS models can help (although space limitations preclude the development of full-blown examples); most of the benefits flow from an increased ability to model subtle but significant agent-environment interactions mediated by situated perception and action. Some balance is introduced by pointing out that this extended explanatory flexibility will often be accompanied by reduced operational transparency (Section 5).

I also suggest (in Sections 5 and 6) that IOS models do not necessarily operate in the mode of incremental revision typical of orthodox OFT, rather they find value in focussing attention on components of foraging behaviour that are either concealed, or exaggerated, through the framework assumptions of orthodox models; for example, IOS models facilitate the generation and evaluation of hypotheses concerning the functional potential of very *simple* mechanisms.

Section 6 further isolates the contributions of IOS models in the context of (i) IBMs in ecology, (ii) ostensibly similar work in SAB/AL, and (iii) the larger debate over the nature of *possible* methodological relationships between SAB/AL and theoretical biology.

2. Optimal Foraging Theory

2.1 Orthodox OFT models

OFT starts from the assumption that foraging behaviour can be treated as a Darwinian adaptation, and, as such, can be considered to be an ‘optimal’ solution to a ‘prob-

lem' posed by (present and/or past) environments. It seeks to understand the foraging behaviour of animals through the identification of what would constitute optimal foraging behaviour given the constraints on the animal and of its (present and/or past) environment. As an example to follow through this discussion, consider the redshank *Tringa totanus*, a shorebird, foraging for both large and small worms (Goss-Custard, 1977).¹

Orthodox OFT models comprise three primary components: a decision variable, a currency, and a set of constraints (Stephens & Krebs, 1986). Their specification does not require that individual foraging agents be explicitly instantiated; orthodox models draw behavioural inferences simply from mathematical (equational) relationships between the components.

All orthodox OFT models consider the 'best' way to make a foraging decision, and the *decision variable* captures the type of choice that the animal is assumed to make (or that natural selection has made for it). In the redshank example, the decision is whether or not to eat small worms upon encounter (it is assumed that large worms will always be eaten), and the variable has two possible states; 'eat' and 'reject'.

The *currency* specifies the quantity that the animal is supposed to be maximising. The ultimate mandate for OFT in Darwinian natural selection requires that this currency be identifiable with reproductive fitness, but since this quantity is more or less impossible to assess in any given study, practitioners of OFT invariably employ a proxy. In the redshank example, the rate of energy intake is the currency, with (it is assumed) large worms providing more energy than small worms.

Constraint assumptions limit and define the relationship between the currency and the decision variable. Intrinsic constraints derive from properties of the animal; the speed with which the redshank can search for prey and the handling time required for each captured worm both qualify. Extrinsic constraints derive from the environment; the densities of the two types of worm serve as extrinsic constraints in the redshank example.

Orthodox OFT models also include a *choice principle* which defines how optimisation of a currency is to be interpreted; this is almost always 'maximisation' but in principle could be some other function (for example a step function would lead to a 'satisficing' formulation of OFT). A *currency function* is also required to transform the decision variable into a value commensurable with the currency. And finally there must be a method of optimisation, ranging from analytic solution to the numerical methods of 'stochastic dynamic programming' which can cope with state-dependent decisions and uncertain environments (Mangel & Clark, 1988).

These components together constitute what may be

called the 'framework assumptions' of orthodox OFT, and are usually defined after a period of observing the behaviour of a focal animal. An equational model is then built, which is used to predict how the animal should behave in different environments in order to be 'optimal', and the extent to which observed and predicted behaviours differ is taken to indicate the degree of revision (in terms of the incorporation of additional components - usually constraints) that the model need undergo; these conclusions are framed as testable hypotheses of the form '(additional) component *P* plays such-and-such a role in observed foraging behaviour *Q*'. This mode of operation may be termed 'incremental complexity' (to use the terminology of Maley, 1999); that is, an initially simple model is augmented with additional constraints until (if ever) it satisfactorily predicts real world behaviour.

Goss-Custard (1977) predicted that the foraging redshank should *either* always accept *or* always reject small worms. He observed, however, that under some conditions redshanks would accept *some* small worms. Revised models, proposed to account for such 'partial preferences', incorporated additional constraints of imperfect discrimination (Rechten et al., 1983) and minimisation of starvation (Houston & McNamara, 1985).

It must be stressed that orthodox OFT models *cannot* tell us whether or not an animal 'optimises'. Some authors have criticised OFT for being 'unscientifically irrefutable' in the sense that 'alternatives to optimality' are never considered, rather, practitioners of OFT insist on "shoring up the cracked facade of optimality with a scaffolding of *ad hoc* modifications" as the imaginary critic of Stephens and Krebs (1986, p.207) has it (for a less lyrical but substantively identical dissension see Pierce & Ollason, 1987). But of course optimality is not the hypothesis under test, it is a way of generating testable hypotheses. In the long run OFT will stand or fall on the knowledge gained from the progressive confirmation or refutation of these hypotheses, and *not* because of any 'demonstration' that animals do, or do not optimise.

2.2 What use orthodox OFT?

OFT models can ask what animals are 'designed' to do, and they can also attempt to unravel the behaviour-generating mechanisms involved.

The first kind of question invokes a functional perspective, and involves finding a set of constraints and a currency (or currencies) such that an observed behavioural pattern may be considered optimal. A focus on currency assesses the validity of a chosen proxy for reproductive fitness; recall the new currency of minimisation of starvation proposed by Houston and McNamara (1985). A focus on constraint addresses aspects of environment structure that have functional significance (extrinsic constraints) - for example the influence of se-

¹A more detailed exegesis of this example can be found in Krebs and Kacelnik (1991).

quential encounters explored by Rechten et al. (1983).

There is a problem here. Given a discrepancy between model prediction and observed behaviour, there is no general rule for deciding which of these foci to attend to, and, worse, it is not even possible to maintain them as distinct options. Attention to extrinsic constraints can be considered to be an elaboration of currency; for example a goal of intake rate may multiply into a set of sub-goals (explore, find food, hoard food, etc.), each associated with (possibly overlapping) sets of extrinsic constraints. On the other hand, recourse to multiple currencies (quantity of food located, quantity of food hoarded, etc.) can represent the interpolation of extrinsic constraints between reproductive fitness and behaviour. The utility of these currencies as proxies for reproductive fitness can then be seen to depend upon the presence of particular extrinsic constraints (food hoarding may only make sense in particular kinds of environment). This conceptual overlap between ‘currency’ and ‘constraint’ is revisited in Section 4.6.

Hypotheses about ‘behaviour-generating mechanisms’ follow from an analysis of decision variables and *intrinsic* constraints, both of which may be considered to represent aspects of the internal mechanism of an animal. For example Cheverton et al. (1985) assume that bumblebees are maximising intake rate and then use ‘errors’ in performance to infer properties of the mechanism controlling movement between flowers; these inferences constitute mechanistic hypotheses. In such instances seemingly suboptimal behaviour can be attributed to the impact of previously unsuspected *intrinsic* constraints.

3. IOS models

3.1 Structure

The use of genetic algorithms (GAs) in the design of situated agents should need little introduction, being prominent both in evolutionary robotics and in the design of software agents throughout the SAB/AL community. For present purposes we shall restrict our attention to the latter, the essential components of which are (i) a computer simulation model of agent and environment framed in terms of situated perception and action, (ii) a mapping from a ‘genotype’ (usually a bit string, or a string of real numbers) to a ‘phenotype’ (the structure, or mechanism of the agent), (iii) a supervening GA which acts over a population of genotypes, and (iv) a fitness function (a metric by which the performance of agents can be evaluated and compared).

Note that it is not intended for any significance be drawn from the ‘evolutionary’ connotations of GAs (although see Section 4.5). Individual-based Optimal Situated (IOS) models can in principle accommodate the use of any optimisation algorithm capable of exploring a space of agent structure and delivering ‘near optimal’

solutions.² The present emphasis on GAs derives both from their successful application to a wide range of optimisation problems, and from their familiarity within the SAB/AL community for the precisely the purposes described above.

3.2 IOS as OFT: At first glance

IOS models can be construed as unorthodox OFT models in the following way. The method of deriving an optimal solution (and the choice principle) in orthodox OFT clearly correspond to the GA in IOS models. Currency corresponds to the fitness function. The constraints are determined by a combination of environment (extrinsic) and phenotype/mechanism (intrinsic) - both those aspects open to variation during optimisation and those that remain fixed. The currency function corresponds to the process of evaluating an agent within its environment and giving it a fitness score. All these equivalences should be obvious. Decision variables, by contrast, do *not* have necessary counterparts in IOS models. It is entirely up to the modeller whether or not the phenotypic architecture utilised mandates interpretation in terms of the operation of a decision variable. Some may do, for example rule-based architectures that hinge on particular internal state variables, and some may not, for example simple neural networks.

To say again, IOS models are nowhere proposed to be ‘better’ than orthodox OFT models. Their relationship is complementary, and a trade-off is already apparent: Whereas orthodox OFT runs the risk of overly restricting the space of possible hypotheses through the biases inherent in the orthodox framework assumptions, IOS models risk confounding *post-hoc* analysis through the complexity engendered by the relatively large number of components and parameters (and interactions between them) constitutive of even simple instances. In the following section we explore reasons for favouring the IOS side of this trade-off, but it should not be forgotten that compelling arguments for engaging in orthodox modelling will always exist, some of which will be outlined in Section 5.

4. How IOS models can help OFT

Orthodox OFT, in spite of (or perhaps because of) its indubitable popularity amongst many biologists, has remained controversial. Some criticisms are frequently rehearsed and futile (c.f. the debate surrounding ‘alternatives to optimality’), of these we will hear no more. Others are well formed and significant, some of these are described and expanded upon below, and still oth-

²‘Near-optimality’, in optimisation theory, describes the results of processes which, although designed to find optimal solutions, cannot be guaranteed to do so; ‘near-optimal’ solutions are (almost) always very good, even if they may not always be the best.

ers are novel to this paper. No comprehensive coverage is pretended here; there is neither the space nor the need. Many criticisms focus on aspects of OFT that are best addressed from within the orthodoxy; see Pierce and Ollason (1987), and for a defence, Stephens and Krebs (1986).

Essential in what follows is the status of *behaviour* as an explanatory entity. The position advocated here is that behaviour is a product of the joint activity of *agent*, *environment*, and (crucially) *observer*; thus (agent-side) mechanisms underlying the generation of any behaviour should not be assumed to be identical to the behaviour itself.

4.1 *Mathematical tractability*

Of those criticisms of OFT that *have* been expressed in the literature, perhaps the most common concern the implausibility of specific constraints often employed in orthodox models for the purposes of mathematical tractability. For example, many conventional foraging models (including the original redshank model of Goss-Custard, 1977) assume (i) the exclusivity of search (for prey) and exploitation (of prey), (ii) sequential pseudo-random encounters with prey, and (iii) that the animal has complete information about the state of the environment. Although the relaxation of these constraints can be, and has been explored from within the orthodox framework (Stephens & Krebs, 1986), IOS models offer avenues for the progressive relaxation of other constraints that would be very difficult to achieve in any other way. For example a constraint of homogeneous within-patch food distribution may resist manipulation in a mathematical model, but is immediately open to relaxation in an IOS model in which items of food may possess explicit spatial locations. There is a continuum of flexibility to be identified here. At one extreme lie the relatively rigid orthodox equational models, in the middle the general class of individual-based models (IBMs) allow assumptions of homogeneous group properties to be modified, and finally IOS models, being situated, permit the relaxation of constraints that embody assumptions about perception and action (recall that IOS models are *defined* as IBMs that incorporate optimality and situated perception and action).

4.2 *Decision variables*

The remaining criticisms concern the framework assumptions of orthodox models in general, and are therefore difficult to resolve from within the orthodoxy itself. The first of these concerns the requirement that internal mechanism be understood in terms of the operation of some kind of decision variable. Although the mathematical instantiation of a decision variable does not imply that it is implemented as such in the animal, in prac-

tice, orthodox OFT provides no way of interrogating this implication. It is then all too easy to read orthodox OFT models as making significant and possibly erroneous claims about mechanism, a position exemplified by Gallistel et al. (1991):

“What is striking about these optimality models is that in spelling out the decision processes underlying the optimisation of foraging behaviour they credit the animal with complex representational and computational abilities.” (1991, p.18)

Also, the distinction between a decision variable and an *intrinsic* constraint does not appear to have any *a priori* justification. Both can be components of internal mechanism, and it may be that optimal behaviour can arise solely from the (externally observed) interplay of constraints (intrinsic and extrinsic) and currencies. There are also the attendant dangers of associating the operation of a decision variable with junctures in behavioural dynamics that may seem significant from the perspective of the external observer but which need not indicate the existence of mechanistic structures dedicated to the entailment of these junctures; as argued above, internal mechanistic correlates for behaviours (or ‘choices’ between behaviours) should not be presumed to exist.

IOS models avoid these difficulties by permitting the construction of models *without* obvious decision variables (for example using simple neural networks). Such models broaden the space of possible internal mechanism, and, in particular, facilitate the generation and evaluation of hypotheses concerning the behavioural potential of very *simple* underlying internal mechanisms. In particular, the risk is no longer run of decision variables being confounded with junctures in behaviour that appear significant from the perspective of an external observer, and by the same token, the awkward (and arbitrary) distinction between decision variable and intrinsic constraint can be abruptly dispensed with.

Consider an example from evolutionary robotics. Mondada et al. (1993) found that it was much easier to evolve a *single* network controller for a complex task (to find and stay close to a target object) than it was to evolve four separate networks corresponding to the separable behavioural elements into which the task could be divided from the point of view of an external observer (these were; explore, avoid walls, approach target, discriminate target from walls). Indeed, they found the latter method impossible, concluding that at least in this case mechanistic and behavioural decompositions were incommensurate. Any attempts to associate (mechanistic) decision variables with the observable junctures in behaviour would therefore have been wholly inappropriate.

4.3 Historical constraints

Another problematic distinction is the separation of constraints into mutually exclusive ‘intrinsic’ and ‘extrinsic’ categories. Such a scheme of classification misses out any constraints that may arise through the history of interaction between agent and environment; these can be called *historical* constraints. One can talk about such constraints at a variety of timescales, notably behavioural, developmental, and evolutionary. The present focus on optimal behaviour encourages concentration on behavioural historical constraints, but evolutionary historical constraints will receive some attention in Section 4.5.³

Consider, as an example, the construction of termite mounds. Some of the essential constraints on the behaviour of an individual termite that lead to it contributing towards the construction of a termite mound can only be framed in terms of the past history of interaction of that termite with its conspecifics and with the surrounding environment (in this instance the phenomenon is also known as ‘stigmergy’).

As an example from the SAB/AL literature, consider the simple model of clustering behaviour in groups of artificial ‘ants’ in Deneubourg et al. (1990). Individual ants wander at random and pick up or put down objects with probabilities determined by the density of objects in the ant’s neighbourhood. Over time more and more locations in the shared environment become ‘blocked’ to the formation of clusters since (if the above probabilities are appropriately selected) no objects are deposited in empty locations; these locations then constrain both the possible locations of clusters, and the fact that clusters will form at all (as they must do as the number of ‘unblocked’ locations diminishes).⁴ Although this behaviour is somewhat simpler than termite mound building, the principle of historicity is the same and indeed the cluster patterns generated by Deneubourg’s ants are very similar to the clustered ‘graveyards’ of ant corpses built by real ants *Messor sancta* (and others).

In general, IOS models can elucidate behavioural historical constraints by providing a sufficiently rich medium in which agent-environment interactions can create dynamical invariants which constrain the future dynamics of the system, and which are able to alter their own conditions of realisation. These are constraints that need not be, and in some cases *cannot* be imposed from the start in either ‘extrinsic’ or ‘intrinsic’ guises. An essential qualification is that such constraints can arise in IBMs at a variety of levels of description. Again there is a continuum; the richer the medium of interaction, the

³The present use of the term ‘historical’ does not need to be yoked to notions of ‘contingency’ or ‘frozen accidents’, as it is to an extent in some discussions of evolutionary history.

⁴The fact that this model does not utilise an optimisation procedure, and is therefore not strictly an IOS model, does not detract from its role in illustrating a behavioural historical constraint.

greater the potential for historicity. IBMs that do not incorporate perception and action may nevertheless permit the emergence of historical processes contingent on the (broadly construed) interaction of distinct agents, but only *situated* models (in the present sense of IOS models) can possibly engender historical processes from a substrate of *sensorimotor* interaction.

4.4 The ‘ecological’ environment

It is difficult, from within the framework of orthodox OFT, to appreciate the structure of the environment from the perspective of the agent itself. And as Pierce and Ollason (1987) surmise:

“It will always be possible to identify a set of environmental characteristics with respect to which observed behaviour is consistent with a particular functional hypothesis, but this process is entirely circular. By asserting that animals perceive the environment in a particular way it would be possible to show that observed foraging behaviour was consistent with any functional hypothesis.” (1987, p.114)

The problem is that orthodox OFT models, in not dealing with explicitly instantiated and situated agents, are framed in terms of the environment as it might appear to us as observers (the ‘external’ environment), whereas many components of such models (notably the decision variable - if there is to be one - and the currency) should properly be framed in terms of the environment as it appears to the agents themselves (the ‘ecological’ environment’).⁵ What is required is a principled strategy for tracing the relationships between external and ecological environments. Orthodox OFT models, again because they do not deal with explicitly instantiated and situated agents, do not provide an effective means of doing this. As Pierce and Ollason conclude, the opportunity then remains for researchers to propose that “animals perceive the environment in a particular way” to suit their functional convenience.

IOS models provide an effective response through allowing the researcher full objective access (in principle) to the *merkwelt* of explicitly instantiated and situated agents. Let me be clear: I do not insist that the researcher *experience* the world as the agent does, merely that she can (in principle) measure all those features of the external environment that constitute the ecological environment of the agent; the levels of activation of the sensors in various conditions, and so on. Thus, the IOS

⁵This environmental terminology is due to Brandon (1990). It is, however, necessary to warn that his own definitions differ slightly from those given here, and he also discusses the separate notion of a ‘selective environment’, being that which is “measured in terms of the relative actualised fitnesses of different genotypes across time and space”.

modeller is *not* free to assert that animals perceive the environment in an arbitrary fashion (although, of course, the structure of the environment itself *is* under the control of the modeller). Instead, the ecological environment of the agent can be constructed by the researcher from the interplay of extrinsic, intrinsic, and historical constraints. Particular functional hypotheses can be entertained only to the extent that they do not conflict with these constraints, and of course it is exactly the nature of these constraints (above all *intrinsic* constraints) that constitute hypotheses about ‘behaviour-generating mechanisms’.

4.5 *The pragmatics of optimisation*

By construing behaviour as an ‘optimal solution’ to an ‘environmental problem’ (Section 2.1), orthodox OFT makes two implicit commitments: (i) that process and outcome (optimal behaviour) are conceptually distinct, and (ii) that evolution is a process of optimisation. These commitments are separate; we will deal with each in turn.

First, and in contrast to orthodox OFT, IOS models can be constructed in which the process and outcome can continuously interact and influence each other. Indeed, as soon as optimisation is *instantiated* as a process there is the potential for this to happen. Such interaction may be particularly significant if it entails the formation of behavioural historical constraints; recall from section 4.3 that such constraints often resist pre-specification as part of a problem for which a solution may be found. In these cases the optimal behaviour at any time (the outcome) cannot be understood without an appreciation of the preceding optimisation dynamics.⁶ However, in order to maintain the methodological parallel between IOS and OFT, it is important to ensure that the GA itself is doing nothing more than optimisation. This means ensuring, as far as possible, that there are no interactions *between* genotypes that could affect the fitness values awarded to the genotypes. In this sense, IOS models retain an independence from the particularities of the method of optimisation.

Second, there are many GA-based models in which the evolutionary (process) mechanics of the GA itself *can and do* directly influence genotype fitness (consider, for example, the co-evolution of pursuit and evasion strategies in Miller & Cliff, 1994). In such cases, neither the outcome nor any process-outcome interactions can be held independent of the process itself.⁷ Such models

⁶Similar issues can arise when dynamic optimisation techniques are applied within an orthodox framework (Mangel & Clark, 1988). However, the considerable potential for behavioural historical constraints in IOS models (Section 4.3) justifies the present emphasis.

⁷Indeed, historical constraints can arise directly from the (process) mechanics of a GA and can influence its dynamics in ways not necessarily associated with the dynamics of any actual behaviours, even if the specification of behaviour is still the outcome (see for

example Di Paolo, 1999, ch. 5). And it is this historical potential - whether behavioural as in Miller and Cliff (1994) or attached to GA mechanics as in Di Paolo (1999) - that partly distinguishes such models from equational game-theoretic models, in which optimisation is frequency dependent but in which optimal solutions (now described as ‘evolutionary stable strategies’) are still considered independent of the mechanics of the (optimisation) process (Maynard Smith, 1982).

exceed the minimum specification of IOS models (laid down in Section 3.1) in that the ‘evolutionary connotations’ of the GA cannot now be discounted, and indeed often become prominent (in the guise of process mechanics) in explanations of model behaviour. At the extreme, GA models can be used to explore pure evolutionary phenomena (see for example Hinton & Nowlan, 1987). These models usually place little if any emphasis on the situated behaviour of individual agents and therefore have little to do with foraging theory, orthodox or otherwise. They are not IOS models.

It is important to note that *even if* the method of optimisation bears explanatory weight, this still does not mean that such models can now test whether or not animals optimise (see Section 2.1). Rather, such models should be interpreted as extensions of IOS models in which aspects of the method of optimisation can constitute constraints on optimisation; such constraints can only modify the *space of generable hypotheses*. This contrasts with both orthodox OFT and ‘unextended’ IOS models (Section 3.1), for which constraints attach only to outcomes or to (optimisation) process-outcome interactions; thereby *generating hypotheses*.

In all cases, the parallel with OFT requires an emphasis on outcome (optimal behaviour) over process. However, whereas orthodox models encourage conceptual separation of process from outcome, IOS models can assess the impact of this separation, in particular through their ability to elucidate behavioural historical constraints, which may interact with the dynamics of optimisation. And ‘extended’ IOS models, whilst less related to orthodox OFT models through their involvement with ‘evolutionary’ mechanics, permit the consideration of a suite of phenomena concerning relationships between optimisation, evolution, and adaptive behaviour. However these are issues beyond the remit of the present discussion, and for what remains of this paper IOS models should be interpreted as ‘unextended’.

4.6 *Proxies for reproductive fitness*

The practice of adopting a proxy for reproductive fitness is widely appreciated as troublesome (Pierce & Olason, 1987, see also Section 2.2). IOS models are at least explicit in directly associating (in the GA) the fitness function (short-term optimality) with reproductive fitness (long-term optimality). There is also the potential for IOS models to employ progressively general fitness functions, ultimately to the extent that no explicit

example Di Paolo, 1999, ch. 5). And it is this historical potential - whether behavioural as in Miller and Cliff (1994) or attached to GA mechanics as in Di Paolo (1999) - that partly distinguishes such models from equational game-theoretic models, in which optimisation is frequency dependent but in which optimal solutions (now described as ‘evolutionary stable strategies’) are still considered independent of the mechanics of the (optimisation) process (Maynard Smith, 1982).

fitness function need be stated at all, with differential reproductive success arising solely out of the interplay of agents and their environments. Thus, at least in principle, IOS models afford a means of evaluating the consequences of making assumptions about how long-term and short-term optimality relate.⁸

5. IOS as OFT: A second look

Thus it is possible to consider IOS models as OFT models in which the constituent assumptions are dramatically restructured. Decision variables become optional, and constraints can be articulated in terms of sensorimotor processes and spatiotemporally defined and structured simulated environments. Ecological environments can be constructed, and the pragmatic value of currencies, and indeed of optimisation itself, can be interrogated.

As with orthodox OFT models, IOS models can be directed towards both functional and mechanistic modes of inquiry, engaging with both well adapted and apparently suboptimal patterns of behaviour. Functional hypotheses for observed patterns of behaviour can be derived from the fitness function, given the nature of the constraints. Insight into mechanistic structures potentially able to underlie observed behaviour patterns can be gleaned from a knowledge of the intrinsic constraints and their (possibly historical) interplay with extrinsic constraints. In particular, IOS models allow for *simple* mechanisms to be considered through (i) abandoning decision variables and (ii) distributing the task of explanation across agent, environment, and time. It should be apparent from all this that IOS models cannot be treated as simple extensions of orthodox OFT models, and some other more profound differences are now discussed.

First, orthodox models are usually of ‘incremental complexity’ (see Section 2). Predictive failures are normally attributed to an inadequate capture of pertinent constraints (or, more radically, currency), prompting a revision of the model. The process ends (if ever) when sufficient constraints have been incorporated such that there is no significant difference between model and real-world behaviour. There is, however, an inconsistency in this approach. The incremental addition of constraints, in orthodox OFT, is always in the context of the significant set of framework assumptions concerning decision variables, extrinsic *versus* intrinsic constraints, and so on. These assumptions are already so strong that it is not clear that *incremental* revision of models in such a context is always going to be appropriate. Of course in many cases progress can be made in just this way,

as the continuing popularity of orthodox methods attests. But it may be, in other cases, that the aspects of model behaviour that prompt revision are in fact consequences of the framework assumptions themselves, and not of inadequate constraint capture. In IOS modelling, these framework assumptions can be relaxed; and so, rather than generating predictive hypotheses of incremental verisimilitude, IOS models can bring into focus aspects of orthodox models that are either explicit but potentially unnecessary (e.g. decision variables) or implicit and usually ignored (e.g. situated perception and action). Unlike orthodox models, IOS models need not operate in a mode of ‘incremental complexity’.

Second, IBMs in general (of which IOS models are a subset) often present ‘explanatory opacity’ (Di Paolo et al., 2000; Grimm, 1999). Even relatively simple IOS models will usually be of sufficiently complex constitution that their operational description is not obvious. In general, work has to be done to understand the behaviour of the model itself. This requires the formulation and assessment of hypotheses - not about the real world - but about how the various structures instantiated by the model entail model behaviour. This may be particularly so, and the process of investigation particularly valuable, when historical constraints are prominent, or when aspects of process mechanics figure in operational descriptions of ‘extended’ IOS models (see Section 4.5). This contrasts with orthodox OFT models which, given sufficient mathematical fluency on the part of the researcher or student, do not hide the means by which model structure entails behaviour.

Third, IBMs (and thus also IOS models) have occasionally been touted as offering greater explicitness and clarity *vis à vis* their assumption structure than mathematical formalisations (Miller, 1995). However, it may equally be argued that IBMs run a *greater* risk of artefactual results than mathematical models because of the large array of components and potentially manipulable parameter settings of which they normally comprise. Assumptions in mathematical models, although they may sometimes be more theoretically pernicious than those of their IBM counterparts, at least have the virtue of being explicitly stated.

A further general criticism of the IOS modelling approach, when held in comparison with orthodox OFT, may be levelled at its apparent ‘exploratory’ character; it may seem that IOS models are best suited to situations in which coherent hypotheses, framed in terms of a well-defined solution space, have *not* been formulated. However this would not be entirely fair. In principle, IOS models are equally amenable to the formulation of explicit hypotheses, it is just that hypotheses particular to IOS models will engage with issues of historical constraints, of sensorimotor interactions (and so on), and as yet there is little precedent for concocting such hy-

⁸Bullock (1999) has advanced a similar claim with regard to the benefits of ‘evolutionary simulation’ modelling, a related methodology that may be distinguished from IOS modelling by (i) an equal emphasis on process and outcome, and (ii) no insistence on modelling situated perception and action.

potheses. Therefore IOS modelling may well appear exploratory (not a bad thing in itself), but this may be a consequence of novelty rather than inherent disposition and ought not to be considered a permanent state of affairs.

6. Related Work

6.1 *Individual-Based Models in ecology*

The previous section began to consider some properties of IOS models (explanatory opacity, a questionable explicitness of assumptions) that inherit from the larger class of IBMs, and some of the benefits offered by IOS models identified earlier also attach, to some extent, to this larger class (historical constraints, and the relaxation of constraints in general). But the fundamental distinguishing features remain: Recent review papers indicate that IBMs in ecology do not generally incorporate optimality (that is to say individual structure is not the result of any optimisation process), nor do they operate at the level of situated perception and action (Grimm, 1999); differing in both regards from IOS models. Consider, as a representative case study, the recent history of modelling the optimal distribution of foragers in patchy environments. Equational models have been applied to this problem ever since the seminal efforts of Fretwell (1972); such models have been reasonably successful in predicting what optimal distributions should look like, but have offered little insight into the nature of the mechanisms by which such distributions might be arrived at. A subsequent wave of IBMs (see for example Moody & Houston, 1995) explore the relationships between individual behaviour and population distribution, but such models have been criticised for not incorporating optimality and thus being inconsistent with the foundations of Fretwell's original intuition. Most recently, Stillman et al. (1997) attempted to model individuals that make 'optimal decisions', but their mechanisms of decision-making are complex and there is still no engagement with situated perception and action. In a related example, Bernstein et al. (1991) analyse an IBM in which agents *are* endowed with a very basic perceptual facility (although individual action is not linked to this in any meaningful way), but in which there is no attempt to incorporate optimality. A fair summary would be that although biologists are well aware of the need to reconcile IBMs with optimality, and also to account for the effects of situated perception and action, principled strategies for achieving these aims are still lacking. IOS models are of course well placed to step into the breach.

6.2 *IBMs in SAB/AL*

Although a good deal of SAB/AL work falls under the IBM rubric, only a small proportion engages with the

same issues as OFT, an example of which is provided by Spier and McFarland (1998) who use situated IBMs to compare the merits of several foraging strategies. However, although their strategies "hav[e] claim to perform optimally when analysed mathematically", their point is exactly that such analysis takes place outside the context of a situated IBM, and they explore how such supposedly 'optimal' strategies fare in situated environments. This work may be taken to illustrate the need for IOS models which, by contrast, allow the exploration of strategies that *are* optimal in the context of situated environments.

Many of the GA-based IBMs that populate the SAB/AL literature are concerned with issues of evolutionary dynamics (Hinton & Nowlan, 1987), and to the extent that they are, their aims should be distinguished from those of IOS models, which are directed squarely at the mechanisms and functions of behaviour. Recall from Section 4.5 that IOS models are characterised by their emphasis on outcome over process and on the independence of outcome from the mechanics of the optimisation process itself. Of course there do exist many examples of models in the SAB/AL canon that *can* be interpreted, in both substance and emphasis, as IOS models (see for example Koza et al., 1992; Seth, 1999). However a detailed survey is beyond the present remit of distinguishing IOS models from ostensibly similar work.

6.3 *SAB/AL and theoretical biology*

Finally, it is important to locate the relationship between IOS and orthodox OFT models in the context of general debate over the nature of *possible* methodological relationships between SAB/AL models and theoretical biology (both broadly construed).

The 'incremental complexity' of orthodox OFT resonates with the 'virtual biology' approach of Kitano et al. (1997) who, although not concerned with 'optimal strategies' of any sort, attempt to introduce as much detail as possible into their models such that there is no (significant) difference between real-world and model behaviour. They are interested in specific biological processes, for example the genetic and biochemical processes underlying the development of the fruit fly *Drosophila Melanogaster*. Their unreserved commitment to verisimilitude encourages them to assert that divergences between model behaviour and actual *Drosophila* development can point to extra factors at play which have yet to be identified by the biologists themselves. Of course the validity of this assertion depends entirely on the theoretical and informational fidelity of their model to the biological state of the art. This approach exemplifies one extreme of the possible relationships between SAB/AL and biology.⁹

⁹Maley (1999) outlines a weaker version of this manifesto, proposing a cycle of activity again similar to OFT in the construction, evaluation, and incremental revision of models. Maley

The other extreme, by tradition associated with Langton (1989), considers that SAB/AL creations provide *instantiations* and not just *models* of biological phenomena, such that observations of their behaviour should be accorded the same status as observations of similar phenomena in the natural world. This ‘strong AL’ position has little contact with the methodological stance of the present enterprise. Perhaps the only direction it provides is in emphasising the utility of analysing behaviours in IOS models that demonstrably do not correspond to observations of the world. However such analysis should be used only to further an operational understanding of the model itself, and should have nothing to do with the postulation of additional empirical data.

Most consonant with the approach as outlined in these pages is the notion of the ‘opaque thought experiment’. Di Paolo et al. (2000) develop the argument that simulation models in SAB/AL provide a way of “re-organising and probing the internal consistency of a theoretical position” such that “theoretical terms may be shown to stand in different relationships than previously thought.” This position does not locate the benefits of SAB/AL in furthering the verisimilitude of models, nor does it interpret these models as ‘instantiations’ of phenomena.¹⁰ Instead, “the researcher may be forced to focus on facts or processes that were at the periphery of her conceptual structure and place them in novel relationships with other theoretical terms” - but only after some work has been done to overcome the ‘explanatory opacity’ of simulation models. IOS models fit in well with this conception, harnessing the versatility of modelling at the level of perception and action to restructure the conceptual apparatus of the researcher with particular regard to the ‘framework assumptions’ of orthodox OFT.

7. Conclusions

As the SAB field continues to mature, it is essential that general methodological positions become elaborated into specific and practical programmes of research. This paper has developed one approach in particular, drawing close parallels between orthodox OFT and the application of optimisation algorithms in situated, individual-based SAB/AL models. This approach highlights the difficulties associated with reconciling the ‘incremental complexity’ ethos of orthodox OFT with the strong

does not insist upon the same degree of fidelity as Kitano et al., a trade-off which affords greater modelling scope but detracts from the confidence with which specific predictions about empirical biology can be formulated.

¹⁰Bedau (1999) has explicated a related idea in which his ‘emergent, computational thought experiments’ are considered to instantiate target phenomena. However, Bedau restricts his interest to ‘deep’ phenomena, for example ‘multi-level emergent activity’, ‘open-ended adaptive evolution’, and ‘unbounded complexity/diversity growth’. Such ‘deep’ phenomena are well insulated from observation in the real world, allowing the distinction between instantiation and model to be somewhat glossed over.

framework assumptions that partially constitute all orthodox models. IOS models allow these framework assumptions to be radically restructured, with decision variables, historical constraints, ecological environments, and the pragmatics of optimisation all receiving attention. These models can be characterised as varieties of ‘thought experiment’, such that any knowledge gained concerns the researcher’s own conceptual space, the re-organisation of which can encourage the formulation of novel empirical hypotheses. Emphasis is placed on the conceptual leverage obtained by modelling subtle agent-environment interactions mediated by perception and action, and in particular, attention is drawn to how IOS models can illuminate the functional potential of relatively simple internal behaviour-generating mechanisms.

From the perspective of ecology, IOS models can be seen as a new generation of IBM in which both optimality and situated perception and action adopt fundamental roles. From the vantage of SAB/AL, IOS models offer a principled strategy for engaging with issues of real currency in the natural sciences.

Acknowledgements

Many thanks to my supervisors P. Husbands and H. Buxton, to the CCNR and the ESPRC for financial support (award no. 96308700), and also to E. Di Paolo, S. Bullock, E. Spier, A. Philippides, M. Quinn, T. Smith, and my reviewers for sound advice.

References

- Bedau, M. (1999). Can unrealistic computer models illuminate theoretical biology?. In Wu, A. (Ed.), *Proceedings of the 1999 Genetic and Evolutionary Computation Conference Workshop Program*, pp. 20–23 San Francisco. Morgan Kaufmann.
- Bernstein, C., Kacelnik, A., & Krebs, J. (1991). Individual decisions and the distribution of predators in a patchy environment II: the influence of travel costs and structure of the environment. *J. Anim. Ecol.*, 60, 205–225.
- Brandon, R. (1990). *Adaptation and Environment*. Princeton University Press.
- Bullock, S. (1999). Jumping to bold conclusions: A review of Amotz and Avishag Zahavi’s *The Handicap Principle: A Missing Piece of Darwin’s Puzzle*. *Adaptive Behavior*, 7(1), 129–134.
- Cheverton, J., Kacelnik, A., & Krebs, J. (1985). Optimal foraging: Constraints and currencies. In Holldobler, B., & Lindauer, M. (Eds.), *Experimental Behavioral Ecology*, pp. 109–126. Fischer-Verlag.
- Deneubourg, J.-L., Goss, S., Franks, N., Sendova-Franks, A., Detrain, C., & Chretien, L. (1990).

- The dynamics of collective sorting: Robot-like ant and Ant-like robots. In Meyer, J., & Wilson, S. (Eds.), *Proceedings of the 1st International Conference on the Simulation of Adaptive Behaviour*, pp. 356–363. MIT Press/Bradford Books.
- Di Paolo, E. (1999). *On the evolutionary and behavioural dynamics of social coordination: models and theoretical aspects*. Ph.D. thesis, University of Sussex.
- Di Paolo, E., Noble, J., & Bullock, S. (2000). Simulation models as opaque thought experiments. To appear in proceedings of Artificial Life VII.
- Fretwell, S. (1972). *Populations in seasonal environments*. Princeton University Press.
- Gallistel, C., Brown, A., Carey, S., Gelman, R., & Keil, F. (1991). Lessons from animal learning for the study of cognitive development. In Carey, S., & Gelman, R. (Eds.), *The epigenesis of mind*, pp. 3–37. Lawrence Erlbaum.
- Goss-Custard, J. (1977). Optimal foraging and size selection of worms by redshank *tringa totanus* in the field. *Anim. Behav.*, *25*, 10–29.
- Grimm, V. (1999). Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future?. *Ecological Modelling*, *115*, 129–148.
- Hinton, G., & Nowlan, S. (1987). How learning can guide evolution. *Complex Systems*, *1*, 495–502.
- Houston, A., & McNamara, J. (1985). The choice of two prey types that minimises the probability of starvation. *Behav. Ecol. Sociobiol.*, *17*, 135–141.
- Kitano, H., Hamahashi, S., Kitazawa, J., Takao, K., & Imai, S. (1997). The virtual biology laboratories: A new approach of computational biology. In Husbands, P., & Harvey, I. (Eds.), *Proceedings of the Fourth European Conference on Artificial Life*, pp. 274–283 Cambridge, MA. MIT Press.
- Koza, J., Rice, J., & Roughgarden, J. (1992). Evolution of food foraging strategies for the caribbean *anolis* lizard using genetic programming. *Adaptive Behavior*, *1*(2), 171–199.
- Krebs, J., & Kacelnik, A. (1991). Decision making. In Krebs, J., & Davies, N. (Eds.), *Behavioural Ecology (3rd edition)*, pp. 105–137. Blackwell Scientific Publishers.
- Langton, C. (1989). Artificial life. In Langton, C. (Ed.), *Proceedings of the Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*, Vol. 6 of *SFI Studies in the Sciences of Complexity*, pp. 1–48. Addison-Wesley.
- Maley, C. (1999). Methodologies for the use of computational models in theoretical biology. In Wu, A. (Ed.), *Proceedings of the 1999 Genetic and Evolutionary Computation Conference Workshop Program*, pp. 16–19 San Francisco. Morgan Kaufmann.
- Mangel, M., & Clark, C. (1988). *Dynamic modeling in behavioral ecology*. Princeton University Press.
- Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge University Press, Cambridge.
- Miller, G. (1995). Artificial life as theoretical biology: How to do real science with computer simulation. CSRP 378. Tech. rep., School of Cognitive and Computing Sciences, Sussex University.
- Miller, G., & Cliff, D. (1994). Protean behavior in dynamic games: Arguments for the co-evolution of pursuit-evasion tactics. In Cliff, D., Husbands, P., Meyer, J., & Wilson, S. (Eds.), *Proceedings of the 3rd International Conference on the Simulation of Adaptive Behaviour*. Cambridge, Mass. MIT Press.
- Mondada, F., Franzi, E., & Ienne, P. (1993). Mobile robot miniaturization: A tool for investigation in control algorithms. In Yoshikawa, T., & Miyazaki, F. (Eds.), *Proceedings of the Third International Symposium on Experimental Robots*, pp. 501–513. Springer-Verlag.
- Moody, A., & Houston, A. (1995). Interference and the ideal free distribution. *Anim. Behav.*, *49*, 1065–1072.
- Pierce, G., & Ollason, J. (1987). Eight reasons why optimal foraging theory is a waste of time. *Oikos*, *49*(1), 111–117.
- Rechten, C., Avery, M., & Stevens, J. (1983). Optimal prey selection: why do great tits show partial preferences?. *Anim. Behav.*, *31*, 576–584.
- Seth, A. (1999). Evolving behavioural choice: an investigation of Herrnstein’s matching law. In Floreano, D., Nicoud, J.-D., & Mondada, F. (Eds.), *Proceedings of the Fifth European Conference on Artificial Life (ECAL99)*, pp. 225–236. Springer-Verlag.
- Spier, E., & McFarland, D. (1998). Possibly optimal decision making under self-sufficiency and autonomy. *Journal of Theoretical Biology*, *189*, 317–331.
- Stephens, D., & Krebs, J. (1986). *Foraging theory*. Monographs in behaviour and ecology. Princeton University Press.
- Stillman, R., Goss-Custard, J., & Caldow, R. (1997). Modelling interference from basic foraging behaviour. *J. Anim. Ecol.*, *66*, 692–703.