Towards a Descriptive Model of Agent Strategy Search

Bruce Edmonds
Centre for Policy Modelling
Faculty of Management and Business
Manchester Metropolitan University
Aytoun Building, Aytoun Street, Manchester, M1 3GH, UK.
Tel. +44 161 247 6479 Fax. +44 161 247 6802
b.edmonds@mmu.ac.uk

Abstract

It is argued that due to the complexity of most economic phenomena, the chances of deriving correct models from a priori principles are small. Instead are more descriptive approach to modelling should be pursued. Agent-based modelling is characterised as a step in this direction. However many agent-based models use off-the-shelf algorithms from computer science without regard to their descriptive accuracy. This paper attempts an agent model that describes the behaviour of subjects reported by Joep Sonnemans as accurately as possible. It takes a structure that is compatible with current thinking cognitive science and explores the nature of the agent processes that then match the behaviour of the subjects. This suggests further modelling improvements and experiments.

1. Introduction

Complex systems are precisely those for which it is extremely difficult to deduce its behaviour from first principles[1]. For example, it is extremely unlikely that one would be able to predict the behaviour of a meercat purely from *a priori* principles, rather one would have to spend a lot of time observing and describing its actions to get a hold on the intricate contingencies of its actual behaviour. With complex systems observation and description must come first and only much later when the detailed behavioural mechanisms are well understood is it sometimes possible to encapsulate some of these in a predictive model. It seems likely that much economic behaviour is complex in this way. This would not be very surprising since it arises as the consequence of the intricate interactions between members of a species that is characterised by the variety and contingency of its behaviour.

But if we are to give up the chimera of numerical predictive models built from *a priori* principles, doesn't that mean we have to give up all formal models and rigour? I would say that we do not. What it does mean, however, is that we have to use formal and computational models that are able to reflect the detailed behaviour as it is observed. We need to constrain our models as much as possible using observations of the relevant phenomena, both in terms of the causal processes as well as the outcomes. Pinning down our models using only the verification of predictive outcomes and an insistence on formal

1. For discussion of the definition of complexity see [2].

simplicity will not be enough. We will need to capture the workings of the processes stage by stage as they are observed *and* reproduce the known outcomes.

In order to perform this feat we will need formal systems that are up to the task of expressing the qualitative cognitive processes that economic processes are rooted in. These more expressive systems come with a price, they are not simple and they seem to allow for multiple representations of the same outcomes. However there is no need for them to be any less formal or rigorous than a set of differential equations.

In this paper I will exhibit an attempt to construct a more descriptive model of the search for an appropriate strategy by the subjects in a particular experiment. It is, of course, impossible to lose *all* assumptions in the construction of any model, but the point is to move towards using fewer and less drastic *a priori* assumptions and use more qualitative and quantitative constraints derived from the processes under study. The purpose of this model is to provide an unambiguous framework for the exploration of the possible processes within these constraints so as to inform the direction of further observation and modelling. This is not merely a static description, for I am not concerned with static phenomena, but a dynamic description of a particular set of observations using the techniques of computational and cognitive modelling. The extent to which this model is generalisable to other phenomena will only become apparent when it is compared with other descriptive models, just as the general characteristics and markings of a species of animal may only become clear when several descriptions of the animals are compared.

To many readers my position will seem too pessimistic. They may be still hoping for some brilliant ‘short-cut’ to a predictive model, that will allow them to miss out the laborious business of observing and describing the underlying processes. However, I would point out that the science of biology has become enormously successful using the methods I am suggesting and, once we have accepted the amount of field work that our subject matter entails, equal success might be achieved in economics.

2. Computational, agent-based models

The move to agent-based models in economics can be seen as part of a transition to a more descriptive style of modelling. An agent-based model must, by its very nature, model a real actor with a computational agent (in some way), so there should be a one-one correspondence between actors and agents. It is not necessary to assume that the law of large numbers will iron out the messy details. The model can allow the global properties to emerge (or not) without having to assume these details away. Real economic actors *are* (almost always) encapsulated, i.e. they will have an inside where the decision making is done which is largely hidden from view, and a series of ways in which they interact with the outside environment which are observable. The agents that are used to model these actors are encapsulated in a directly analogous way[1].

However, many agent-based modellers do not see the need for any greater descriptive accuracy than this. Thus when inspecting the learning, inference and decision making processes that an agent uses in such a model, one often finds something as

1. The fact that agent-based models are typically simulated on a computer forces another descriptive aspect upon the modeller – the decision making processes inside an agent must be computationally feasible.

unrealistic as a simulated annealing algorithm or standard genetic algorithm. These are algorithms that have been taken from the field of computer science, regardless of their descriptive appropriateness for the actual economic actors being modelled. Now it is possible that in some circumstances such algorithms will give acceptable results for the purposes of some models, but at the moment we can only guess whether this is the case. It is not only that we do not know the exact conditions of application of each algorithm, we do not know of even a *single* real circumstance where we could completely rely on *any* of these 'off-the-shelf' algorithms to give a *reasonable* fit.

To be clear, I am not criticising looking to computer science for ideas, structures and frameworks that might be used in modelling. Being a bounteous source of possible types of process is one of the field's great contributions to knowledge. What I am criticising is the use of such algorithms without either any justification of their appropriateness or modification to make them appropriate.

Thus many agent-based models fail to escape the problems of more traditional models. They attempt to use some ensemble of interacting agents to reproduce some global outcome without knowing if the behaviour of the individual agents is at all realistic. The wish for the 'magic' short-cut is still there.

Clearly what is needed is some way of modelling the behaviour of economic actors by computational agents in a credible way. As noted above, real economic actors are probably complex in the sense that it is unlikely that we will be able to deduce their actions from *a priori* principles. What we can do is to constrain our models *as much as possible* from what we know. There are several sources of such knowledge.

1. We can ensure that the global outcomes of the model match the global outcomes of real actors in the standard way.

This is a good start, but when one is using a more expressive formal system like a computational one then this is unlikely to sufficiently constrain the possible models. In other words, there are likely to be many computational model which produce the same global outcomes.

2. We can ensure that the actions of the *individual* actors match those of our agents' behaviour *as* they learn and interact.

Axtell and Epstein set out some criteria for the performance of multi-agent simulations in [1] In this: *level 0* is when a model caricatures reality at the global level through the use of simple graphical devices (e.g. animations or pictures); *level 1* is when the model in is qualitative agreement at the global level with empirical macro-structures; *level 2* is when the model produces qualitative agreement at the agent level with empirical micro-structures; and *level 3* is when the model exhibits quantitative agreement at the agent level with empirical microstructures. The constraint I am suggesting corresponds to their '*level 3*' with an emphasis on the agreement over time.

3. We can look to the emerging guidelines coming from cognitive science as to the sort of learning and decision processes humans might use.

Now the task of the cognitive scientist is difficult, but such scientists are able, at least, to exclude some mechanisms for explaining behaviour and make suggestions for the mechanisms derived from a lot of observation. It is notable that many successful sciences take their ultimate grounding for the behaviour of their components from outside their discipline – chemistry is grounded in physics and biology is grounded in chemistry.

4. We could simply ask the actual actors why they made the decisions they did and how they learnt what they learnt. This is a particular example of using the techniques of business history to extract information demonstrating the way institutions and individuals did behave.

This method has its known drawbacks, but can be successfully used, especially when confirmed by other methods. In any case it is likely to produce more useful and accurate information about the real behaviour of actors than is implicit in many commonly used assumptions.

3. The target behaviour

The behaviour I am aiming to capture in my descriptive model is that specified by the results of 'experiment 1' described by Sonnemans in [12]. This is an experiment where subjects have to repeatedly sell a notional item in a bidding process. They receive a sequence of offers for this item. Each offer costs them 2 cents and each offer is drawn randomly from the interval [1,100] – they are paid the amount of the offer that they accept minus the costs incurred (2 cents per offer accepted) minus a fixed fee of 50 cents per game. Thus they have opposing incentives: to wait for a better offer and to avoid excessive cost. The 'optimal' strategy is to have a reserve price, that is wait for a offer of 81 or over and then accept it.

The experiment is divided into parts: 0, 1, 2, 3, and 4, each of 4, 15, 15, 15 and 20 games respectively. Part 0 is the practice stage where the agent learns but statistics are not kept and no earnings gained. In parts 1, 2 and 3 the agents learn and earn as they do so. In part 1 the game starts in earnest, so that the subjects earn real money dependent upon their performance. At any stage in the game the subjects have the option of finding out any combination of the following information about the game so far: the number of bids; the last bid; the highest bid so far; the cost of bids if they stopped; and the earnings if they stopped. Part 2 is the same as part 1, except that at any stage the subjects can only access *one* of the above pieces of information (this does not stop them remembering or working out this information in their head, of course). In part 3 the first 0, 1, 2, 3 or 4 offers (determined randomly) were automatically accepted, the subject deciding when to stop after that (these offers still had to be paid for). In the last part (part 4) a constant strategy (the best learnt by the end of part 3) is kept for each game, so no learning occurs but statistics are kept and payments made.

What makes this particular experiment appropriate for this purpose is that Sonnemans extracts the strategies that the subjects end up with in a form that is computationally modellable. In the experiment there were several parts. In the initial parts the subjects were able to try out their strategies. Before the last part they had to specify the strategy that would determine offer acceptance for the final 20 games. In all but two cases these could be formulated in terms of five predicates and two Boolean operators. This is not surprising as Sonnemans had done a pilot study to determine the operators that most people would use. These were: $H \geq x$ (the highest offer so far is not less than x); $L \geq x$ (the last offer was not less than x); $N \geq x$ (there have been x offers or greater); $E \geq x$ (earnings are not less than x); $O \geq x$ (there have been at least x offers in a row since the last highest offer or more); AND (boolean conjunction); and OR (boolean disjunction). Thus the strategy 'Accept the highest offer if my earnings are at least 70 or there have been 10 offers' could be expressed as 'H≥70 OR N≥10'.

Two of the strategies were of a form which would entail a non-zero probability of never terminating (e.g. only stop if earnings are greater than 90). Two of the strategies were of the form stop if the earnings are at least X(t), where X is a function and t is the earnings so far. These four strategies were excluded from the results reported by Sonnemans[1]. The other 34 strategies are shown in Table 1.

Subject	Strategy	Earnings Mean	SD	Stopping Decision % Same as actual part 1	% Same as actual part 2	% Same as optimal
2	L>=80 OR H>=80 OR N>=10 OR E>=30	29.7	11.8	53	67	86
4	L>=90 OR (H>=80 AND N>=10) OR E>=30	30.0	11.5	47	73	76
5	(L>=65 AND E>=6) OR H>=70 OR N>=15 OR O>=5	26.6	12.1	80	87	55
6	H>=80	30.5	10.4	47	87	95
7	L>=80 OR (H>=66 AND N>=8) OR E>=20 OR O>=5	28.5	12.1	47	67	69
8	H>=80	30.5	10.4	80	87	95
9	L>=75	29.8	10.0	67	100	77
12	L>=80 OR E>=20	29.8	9.9	53	93	81
13	L>=100 OR H>=70 OR N>=10	28.2	11.3	47	53	63
14	H>=84 OR E>=25	30.3	10.9	80	80	88
15	L>=80 OR N>=10 AND E>=0)	30.1	10.8	80	100	88
16	H>=70 OE R>=10	26.9	11.1	80	67	56
17	L>=80 OR E>=24 OR O>=8	30.0	10.9	73	87	86
18	L>=72 OR H>=75 OR N>=15 OR E>=20	29.1	10.4	47	73	68
19	(H>=80 AND E>=20) OR (N>=10 AND O>=8)	29.1	13.8	60	73	84
20	L>=70 OR H>=70	28.6	10.4	67	87	64
21	H>=76 OR E>=18	29.3	9.9	80	93	72
22	N>=21 OR E>=20	29.1	11.7	53	67	73
23	H>=70 OE R>=10	28.6	10.4	47	53	64
24	L>=82	30.5	10.9	80	80	95
25	H>=75 OR N>=10 OE E>=15	28.1	11.2	67	87	63
26	(H>=85 AND E>=18) OR N>=10	29.5	12.5	40	67	72
27	H>=70 OR N>=15 OR E>=10	26.9	11.2	47	53	55
28	(H>=75 OR N>=6) AND E>=15	29.8	10.2	47	73	76
29	(N>=6 AND H>=80) OR E>=15	28.7	10.1	67	53	68
30	H>=85 OR N>=8 OR E>=20	28.5	12.2	60	73	69
31	L>=75 OR N>=10 OR (H>=50 AND O>=5)	28.7	11.9	60	73	70
32	L>=82 OR N>=9	29.4	12.4	67	73	82
33	N>=7 OR E>=15	27.1	12.7	67	67	59
34	L>=70 OR (N>=5 AND E>=10)	28.6	10.4	60	80	64
35	N>=10 OR E>=10	26.6	11.7	80	73	55
36	H>=68	28.0	10.7	73	87	60
Av		**28.9**	**11.2**	**62.6**	**76.0**	**72.8**
SD		***1.2***	***0.9***	***13.3***	***13.1***	***12.3***

Table 1: Derived from Table 3 of [12], page 317. The strategies of the subjects in part 4. The characters in the second column have the meanings as describe above. The earnings statistics and the percentage same as optimal are based upon a simulation, N=160 000.

As has been remarked in other experimental studies (e.g. [4, 6, 11]): subjects learn better strategies as they gain experience; overall efficiency is reasonably high, and there is a marked tendency to stop too early. This last fact is shown by the stopping statistics for the experiment shown in Table 2.

1. It is not entirely clear why these were excluded from the listing. Presumably because their optimality could not be calculated since they allow for the possibility of unlimited losses.

	Mean number of bids		Stopping			Mean earnings		
	Optimal	Actual	Early	Optimal	Late	Optimal	Actual	Efficiency
Part 1	4.73	4.07	40%	54%	6%	29.4	24.7	84%
Part 2	5.53	4.62	17%	74%	9%	27.3	25.8	95%
Part 3	5.67	4.55	30%	61%	9%	29.9	26.8	90%

Table 2: Derived from Table 1 of [12], page 316. Search behaviour, 36 subjects, 15 periods per part. The earnings are in cent per period.

As Sonnemans points out, this early stopping behaviour is not explained by simple risk aversion in many cases. Figure 1 shows a plot of the mean earnings against the spread of earnings for the 34 strategies listed in Table 1. The line shows the orbit of optimal strategies for various risk/expected earnings trade-offs. Many of the strategies chosen by the subjects were under the top part of the line, so that there were other strategies they could have adopted with the same level of risk but greater average earnings.

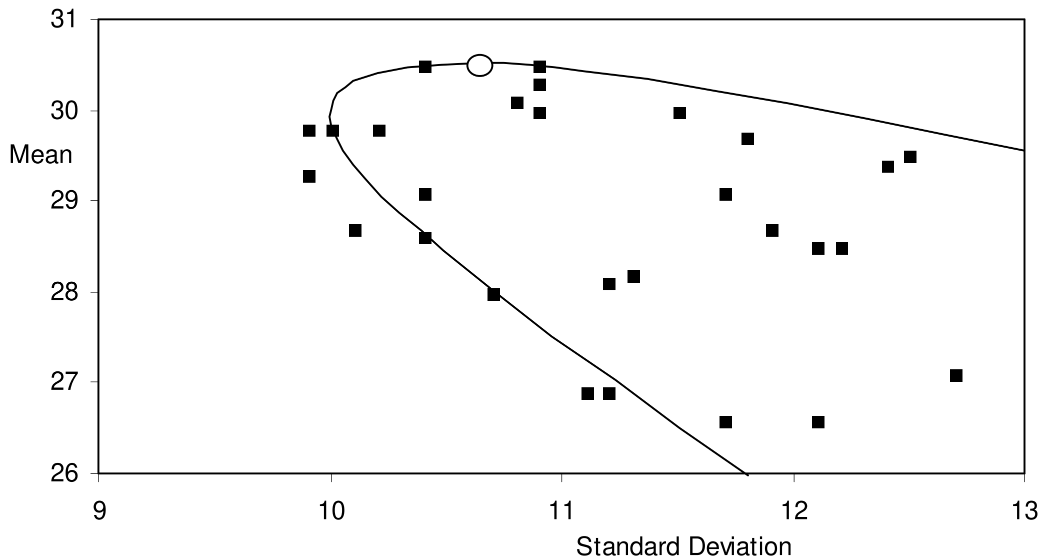


Figure 1: Derived from Fig. 1 of [12], page 319. The curve is the expected mean and standard deviation of the earnings of the strategies of the 'optimal' form (i.e. those of form 'H≥*x*') and the black squares are the strategies that the subjects specified for part 4 (see Table 1). The expected performance by the risk-neutral 'optimal' strategy (i.e. H≥81) is indicated by the O (SD=10.6, Mean=30.5).

4. Hints from cognitive science

What makes Sonnemans' experiment particularly interesting is that it allowed the subjects reasonable scope within which to express their strategies. Even though the scope was still fairly restrictive (as compared with natural language), the subjects specified a range of different kinds of strategies. Clearly, the subjects are learning the *form* of their strategy as well as its parameterisation. This fits into the formal framework for learning proposed in [9]. He also has collect some statistics about their behaviour at different stages of the experiment (Table 2). However these are not sufficient to constrain the possible learning mechanisms employed by the subjects (it was, of course, not designed with this in mind[1]). Thus I have had to look elsewhere for additional constraints – in this case, cognitive science.

I am not suggesting that economics should (or could) be reduced to cognitive science, just that we should accept constraints about the *behaviour* of our units from it – just as chemistry ensures that its models are consistent with what physics tells them about the behaviour of atoms. Nor am I suggesting that we should invade the field of cognitive science – our job is to ensure that the behaviour of our agents matches that of real economic actors so that we can safely explore what happens when they interact monetarily. The job of investigating the purely internal cognitive processes is best left to cognitive scientists. However, just as good chemists must know enough about the physics of atoms in order to ensure that their chemical models are consistent with the laws of physics so must we inform ourselves sufficiently about cognitive science in order to ensure that our agents are compatible with what is known about the workings of humans.

My source is ‘Induction’ by Holland, Holyoak, Nisbett and Thagard [5]. This is a synthesis of the relevant cognitive science and associated philosophy on how induction does and can occur, written from a computational point of view. This attempts to cover all the different aspects of induction (e.g. induction in science), so I will only draw on the basic framework that it suggests. In particular, I have abstracted the following aspects:

- That induction works on a population of ‘rule-based mental models’;
- That there are operations that produce new models from the existing ones;
- That there is some ‘credit assignment’ mechanism that allows the hypotheses to be evaluated in response to their success.

Although [5] uses the paradigm of the ‘classifier system’ as a concrete framework to explore and illustrate the possibilities, I will use the Genetic Programming (GP) paradigm [7][2]. This provides a flexible framework to implemented this picture of induction in an artificial agent. It is more appropriate for this task because the form of the strategies that Sonnemans reported can be directly implemented using GP-based techniques. The GP paradigm involves a collection of expressions that conform to a formal grammar, which are evolved by the operations of combination, variation, propagation and selection, in response to the tasks it is presented with. The original GP *algorithm* was designed for efficient machine induction, so there is no reason, a priori, to suppose that it is a good picture of human induction. However the *paradigm* is much wider, providing a framework within which there are many possible algorithms. This flexibility enables us to search for a descriptive fit with the observed results. The use of this paradigm is examined in more detail elsewhere [3].

1. It is unfortunate that I have not had access to the game-by-game records of the subject’s choices.
2. Strictly, since the expressions are typed (Boolean and numeric) this is an example of strongly-typed GP [8].

5. The model

5.1 General simulation structure

The structure of the computational model follows that of the Sonnemans experiment as closely as possible. Thus there are 36 experiments in parallel, each with one agent. The agents have to decide when to stop a sequence of offers and accept the value of the highest offer so far (minus the cost of the offers and a fixed 'fee' of 50 cents). Thus in each run of this simulation each of 36 agents plays 69 games.

Thus each agent has to decide for each offer in each game whether to ask for another offer or accept the highest offer so far and at the end of part 3 to specify a fixed strategy (of the form discussed previously) for part 4. The material it has to work on is its experience in terms of its earnings and costs in previous games. How the agents specify strategy and make decisions based on this experience is the nub of the model.

5.2 Agent model

Following the hints in Section 4, each agent has a small population of strategies which represent its current hypotheses. These are in the form of those in Table 1. As a result of the experience of each game the agent evaluates its hypotheses, keeps some, changes some, combines some and forgets the rest. In any particular game it uses the strategy that it has evaluated as its best to decide when to stop the bidding process. At the start each agent is given an initial population of these strategies generated at random, and from then on the agent works with these to produce newer and better strategies, so the population evolves and the strategies improve.

There are four key questions to be explored within this framework. Each question concerns how the agent model should be structured so that it reflects (as far as possible) the behaviour of the subjects in Sonneman's experiment. For each question I list some possible answers and explain how the model can be adapted to implement these.

1. Exactly how should the agents adapt their existing hypotheses in the light of their experience?

There are many possible operations that a subject could use to produce the next population of hypotheses. I list the selection that I implemented.

election. This is the operation of keeping the best existing hypothesis, as currently evaluated. The idea is that one holds on to one's best hypothesis.

propagation. Select a hypothesis probabilistically according to its current evaluation and keep it. Since selection is probabilistic this means that you keep a selection of your current hypotheses biased towards those that are currently doing best, but you might sometimes keep less good hypotheses as well.

new. Introduce a totally new randomly generated hypothesis. If one does not introduce any new hypotheses there is a danger of stagnation after a while.

generalisation. Select two existing hypotheses probabilistically according to their current evaluations and join them into one new hypothesis using the boolean 'OR' function. Thus if the selected hypotheses were '$H \geq 70$' and '$N \geq 7$', the new hypothesis would be '$H \geq 70$ OR $N \geq 7$'.

specialisation. Select two existing hypotheses probabilistically according to their current evaluations and join them into one new hypothesis using the boolean 'AND' function. Thus if the selected hypotheses were 'H≥70' and 'N≥7', the new hypothesis would be 'H≥70 AND N≥7'.

join. Select two existing hypotheses probabilistically according to their current evaluations, randomly choose an appropriate function and join them into one new hypothesis using this function. The 'generalisation' and 'specialisation' operations described immediately above are special cases of this. This could be seen as a general exploratory constructive operation.

cut. Randomly choose a node in an existing hypothesis (chosen probabilistically according to its current evaluation) and copy and keep the subexpression as a hypothesis. This might be associated with an attempt at simplification of a more complex expression.

graft. Select two existing hypotheses probabilistically according to their current evaluations, randomly choose a node in one of them graft the other to that node replacing any subexpression that was there before, keep the result. This is another exploratory constructive operation attempting to utilise the properties of simpler hypotheses to make a more effective one.

cut and graft. Select two existing hypotheses probabilistically according to their current evaluations, 'cut' a subtree from the first (as above) and 'graft' it into the second (as immediately above), keep the result. This is similar to the 'graft' operation immediately above, but does not have the disadvantage of always increasing the size and depth of the expressions.

crossover. Select two existing hypotheses probabilistically according to their current evaluations, randomly choose a subnode of each expressions and swap the subexpressions these indicate between the two hypotheses to make two new hypotheses, keep both. This is rather like two 'cut and graft' operations combined. This operation is used very effectively techniques of automatic induction because it is exactly conserves the nodes in the expressions while shuffling around their configurations.

hillclimb. Select an existing hypothesis probabilistically according to its current evaluation, randomly pick a numeric parameter, increment or decrement that parameter, and keep the result. This is a classic exploratory move allowing the incremental parameterisation of an expression.

Thus to take some example mixes of these operations. A mix of 90% crossover and 10% propagation would implement a classic genetic programming algorithm. A mix of 10% election and 90% new would implement a search which merely tries out expressions at random. A mix of 50% hillclimb, and 50% propagation would start with a randomly generated set of expressions and then incrementally explore parameterisation of these.

In the agent model I use here, the proportion of the operations used by agents is set by the programmer.

2. In what way should the agents evaluate their hypotheses?

Clearly the most obvious way in which the expressions might be evaluated is according to the net earnings that using each hypothesis would have generated on the offer sequences that are encountered. However that is not the only possibility. The agent could be also biased in favour of syntactically simpler hypotheses or be in favour of those strategies that incur fewest costs.

As with the operation of variation there are quite a number of ways of combining these factors in the evaluation of hypotheses. It is possible that the subjects are attempting to maximise some function of these, but it is also possible that these factors might be applied separately, for example by eliminating all those which perform badly in any of these respects.

In this model each of these three aspects (earnings, costs and syntactic complexity) are evaluated for each strategy. There are then several different ways of treating the evaluation for each aspect. I chose four: a set proportion of the worst strategies with respect to each can be discarded and the rest just kept on; the strategies would have to reach a minimum level in order to be kept on; the strategies can be given a simple score with respect to each aspect; or they can be ranked.

3. How far back should the agents evaluate new hypotheses against past experience?

An optimal agent would remember the offer sequence in every single past game and use all of them to evaluate its strategies. This would be implausible in this case as the subjects were not allowed to keep notes and would have had to rely on their memory. Thus there is the question of how far back subjects recall in the process of evalutating their hypotheses. Broadly, the further back a subject recalls the more efficient and less contingent is their strategy development.

4. What number of hypotheses should the agents hold at any one time?

It seems clear that human beings would not merely consider one strategy for each game, starting completely afresh if this one was unsuccessful. Rather it is likely that they would recall previous strategies they had tried or thought of as a basis for further strategies. Thus, in effect, the subjects will have access to a variety of past strategies at any stage. However it is also clear that the ability of human subjects to recall (or reconstruct) past strategies is limited, in other words we forget some of our past thinking. Thus what would be a realistic number of strategies for an agent to hold within this framework is also a parameter to explore. Broadly, large populations implement a more exploratory search while small populations implement a more ‘path-dependent’, contingent search.

Clearly, the total number of possibilities that could be constructed by varying the above parameters is vast. However, thanks to Sonnemans’ data, we are able to pin it down in several different ways. What I will do is to exhibit the first steps in this modelling process. This is a process that obviously will need to be continued.

The model was implemented in the language SDML [10]. The code for the simulations described here will be available at: http://www.cpm.mmu.ac.uk/~bruce/tdmass/code.html.

6. The results

Many different simulations were run. Each run was set up with a different combination of operators, model parameters and evaluation methods. There is not room to display all of these here but I will summarise the results and show some of the more significant simulations in more detail. The full list of runs with their set-ups and results will be made available at: http://www.cpm.mmu.ac.uk/~bruce/tdmass/results.html.

For ease of reference the two 'signature' graphs of the target behaviour are shown in Figure 2 below. The first is a rescaled version of Figure 1 and the second is a barchart showing the average proportion of times that subjects earlier than the optimal, at the optimal point and after the optimal point for parts 1, 2 and 3. Please note that whereas Sonnemans' calculated the expected average earnings and spread of earnings using 160,000 simulations of randomly generated games, I had only the computational resources to do relatively few (400 simulations). Thus while the diagram in [2] are accurate to one decimal place (with a probability of 95%), mine are only accurate to the nearest 1 (in the mean earnings), and will overestimate the expected spread of the earnings.

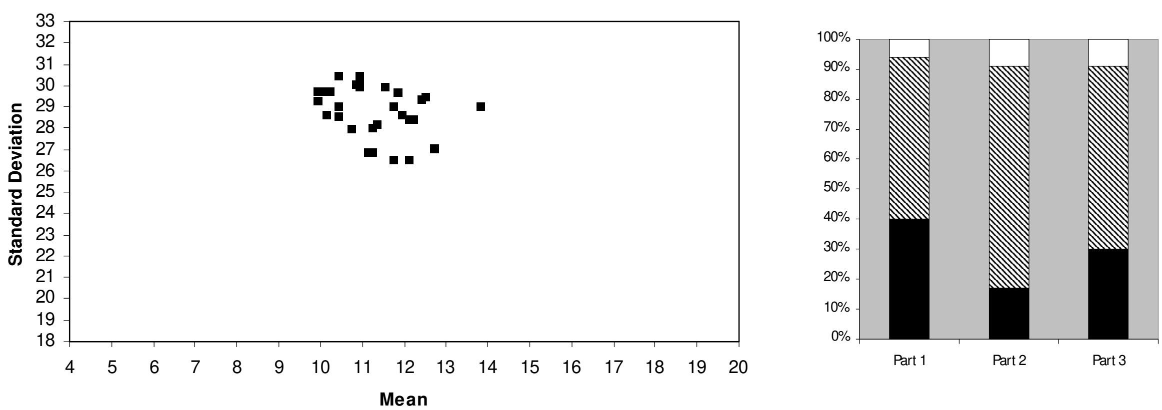
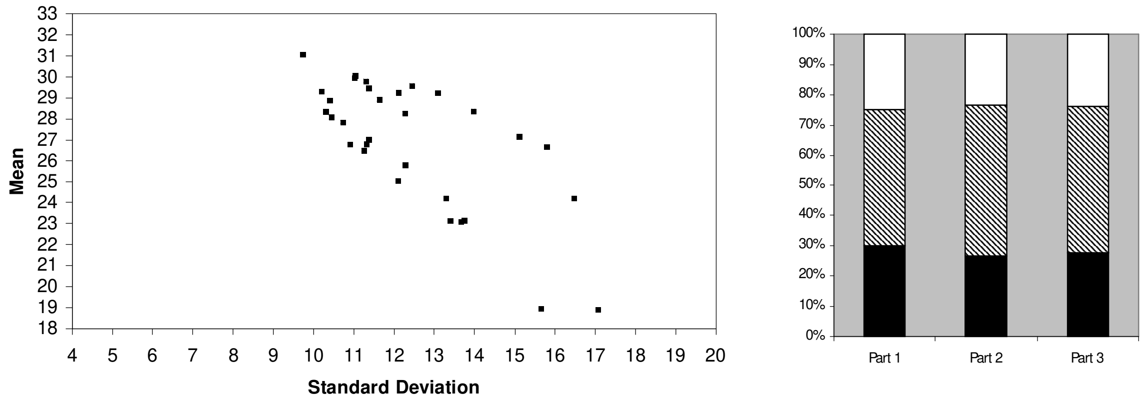
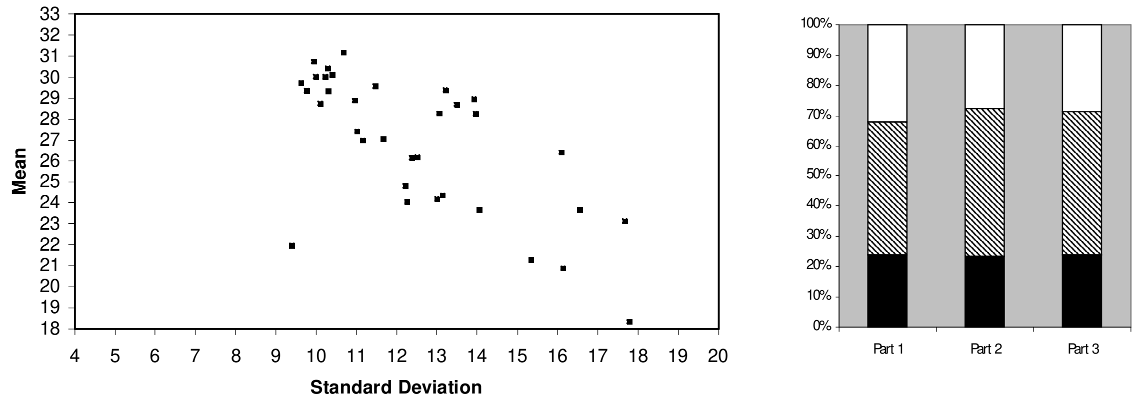


Figure 2: Two 'signature' graphs of Sonnemans' results [12]. The first is a plot of the standard deviation of net earnings against the mean of those earnings over 160,000 random games for the 32 strategies shown in Table 1. The second is the proportion of early, optimal and late stoppings for the 32 subjects for parts 1, 2 and 3, taken from Table 2. The black are is the early stopping proportion, the shaded the optimal and the white is the late.

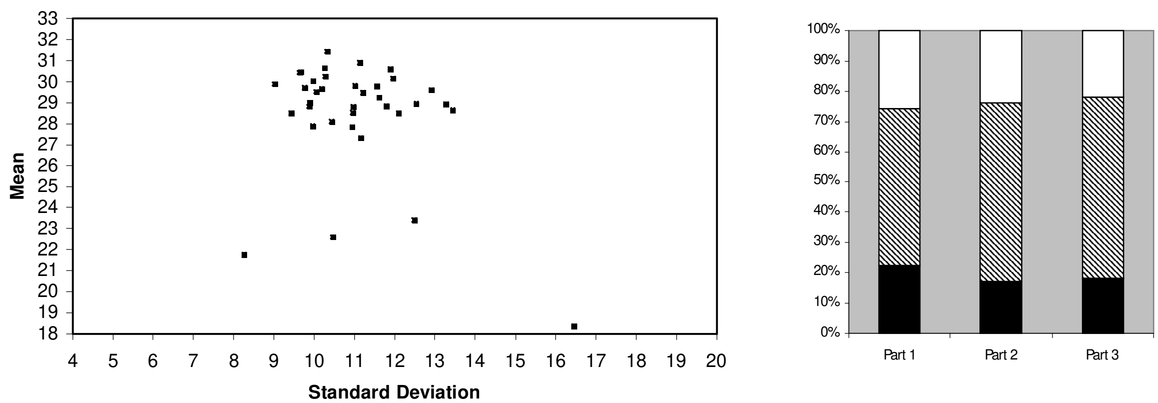
The first three runs mimic the original GP algorithm (i.e. as described in [7]). That is, the proportion of operators was: 90% crossover and 10% propagation; evaluation of strategies was done primarily on net earnings but with a significant bias against syntactically complex strategies; the strategies were allocated a final fitness depending on their score in the evaluation; the initial population of strategies was generated with a maximum depth of 3; and the strategies were evaluated over the previous two games each time. The three runs are with 10, 20 and 30 strategies respectively. The corresponding graphs are shown in below.



GP run with 10 models



GP run with 20 models



GP run with 30 models

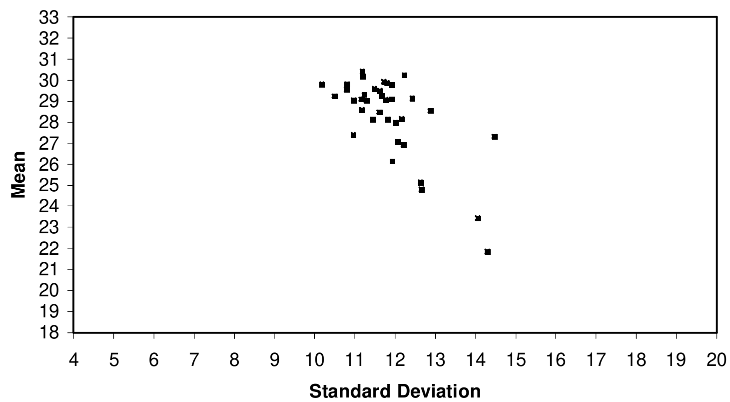
Figure 3: The ‘signature’ graphs of the three GP runs[1].

As we can see from these results although the mean and spread of the earnings for the run with 30 models is approximately right (except that it is bit too evenly spread and there are a few stragglers with low mean) the stopping pattern is not. Unsurprisingly the

1. Note that for the simulation results the mean and standard deviations of the earnings of the final strategies were only estimated over 400 random games.

GP algorithm is not biased towards stopping early. Also the set of strategies the algorithm finished with were not as complex as those of the experimental subject, in fact they were all of zero depth (i.e. involved only one predicate and no Boolean operators), they were: H>=68, L>=68, H>=86, L>=76, H>=84, L>=81, H>=78, E>=69, N>=10, L>=74, E>=74, H>=76, H>=77, E>=63, H>=82, E>=62, L>=74, E>=71, H>=78, E>=75, L>=79, H>=80, E>=71, E>=70, E>=41, H>=76, O>=4, E>=72, H>=74, H>=81, E>=69, H>=75, H>=88, H>=82, E>=54, and H>=75. It seems that this version of strategy search (unlike Sonnemans' subjects) quickly settles on search the combinatronic variations of numbers and single predicates. This explains the few 'stragglers', who just happened not to have an appropriate combination in their initial population of expressions; there is no novelty or hill climbing in this run so they could never develop them.

I do not have enough space to describe all the runs, so I will now turn to the run that most closely matched Sonnemans' results. This was the best run I have achieved in terms of the level of qualitative fit. In this run: the agents had a population of 80 models generated with an initial depth of 1; the proportion of operators was 10% propagation, 2% election, 15% generalisation, 30% hill climb, 4% new, 5% specialisation, 20% cut and graft; the strategies were evaluated over the last 4 time periods in proportion to net earnings only but strategies deeper than and those that always incurred the maximum amount of cost were always removed.



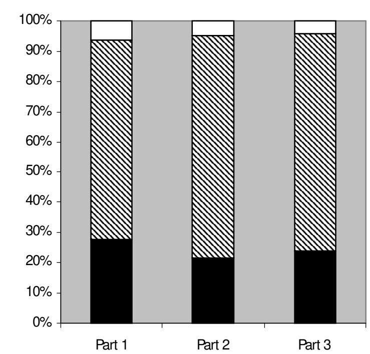


Figure 4: The run wit the most realistic search pattern

This run was closer on the stopping pattern but not completely realistic in the pattern of earnings and their spread – the more successful agent look realistic but there is a 'tail' of unrealistically poor strategies with low average earnings. It did produce a more realistic mix of strategies in terms of a mix of depths.

The equivalent of Table 2 for this run is shown below in Table 3. The stopping pattern is similar except for the fact that for Sonnemans' subjects they exhibited the most optimal stopping behaviour in part 2 rather than part 3 as below. This may have been because in the original experiment the restricted information available to the subjects may have encouraged them not to recall previous highest bids but concentrate on the immediate one and thus improve their performance. This improvement in the absence of recall has been noted in previous experiments (e.g. [4]). the efficiencies of the subjects behaviour shows a similar pattern and can be attributed to the same reason.

	Mean number of bids		Stopping			Mean earnings		
	Optimal	Actual	Early	Optimal	Late	Optimal	Actual	Efficiency
Part 1	4.62	3.60	28%	66%	6%	29.4	27.0	92%
Part 2	4.27	3.68	21%	74%	5%	30.8	28.7	93%
Part 3	5.43	4.73	24%	72%	4%	28.0	26.0	93%

Table 3: The equivalent of Table 2 above for the best simulation run.

The equivalent of Table 1 is shown below for this simulation run (Table 4). Here we do see a variety of simple and complex strategies, and the average and standard deviation of the earnings achieved was close to that found by Sonnemans. What we do not see is the significant presence of strategies using primitives of the form N≥*x* or O≥*x*. It is likely that the bias against excessive cost might have the effect of removing these at an early stage.

Subject	Strategy	Earnings		Stopping	Decision	
		Mean	SD	% Same as actual part 1	% Same as actual part 2	% Same as optimal
1	[AND ['E>=' [80]] ['E>=' [65]]]	28.6	12.9	53	73	72
2	['H>=' [67]]	28.6	11.2	60	67	62
3	['H>=' [72]]	29.1	11.2	67	80	71
4	['H>=' [78]]	29.8	10.8	100	80	90
5	[AND ['E>=' [73]] ['E>=' [57]]]	29.5	11.6	87	73	84
6	['H>=' [77]]	30.4	11.2	73	73	87
7	['E>=' [70]]	29.3	11.7	73	73	79
8	['H>=' [74]]	29.6	11.5	60	80	78
9	[OR ['H>=' [78]] ['H>=' [89]]]	29.1	11.9	67	73	88
10	[AND ['E>=' [56]] ['E>=' [66]]]	28.1	11.8	67	87	74
11	[OR ['N>=' [6]] ['H>=' [79]]]	27.3	14.5	40	60	72
12	['H>=' [68]]	28.0	12.0	80	67	63
13	['H>=' [78]]	29.0	11.3	73	73	87
14	['E>=' [72]]	29.1	11.0	93	93	83
15	[AND ['H>=' [75]] ['H>=' [58]]]	29.9	11.8	67	93	79
16	['H>=' [73]]	29.3	10.5	93	80	74
17	['H>=' [61]]	25.2	12.6	53	73	49
18	['H>=' [79]]	29.9	11.7	73	87	92
19	[AND ['H>=' [56]] ['H>=' [78]]]	29.8	11.9	93	80	91
20	[OR ['H>=' [52]] ['H>=' [63]]]	21.9	14.3	53	47	40
21	['E>=' [70]]	29.3	11.2	73	67	81
22	['H>=' [82]]	30.2	11.2	47	73	96
23	['E>=' [55]]	24.8	12.7	73	67	51
24	[OR ['E>=' [69]] ['E>=' [70]]]	28.5	11.6	87	100	77
25	['E>=' [73]]	30.3	12.2	53	67	86
26	['H>=' [65]]	26.9	12.2	47	67	58
27	[AND ['H>=' [55]] ['H>=' [54]]]	23.4	14.1	40	33	47
28	['E>=' [62]]	27.4	11.0	80	80	64
29	[AND ['H>=' [83]] ['H>=' [83]]]	29.2	12.4	80	80	92
30	[AND ['E>=' [74]] ['E>=' [74]]]	29.1	11.8	53	73	81
31	['H>=' [77]]	28.2	12.2	20	73	83
32	[OR ['H>=' [76]] ['H>=' [67]]]	27.1	12.1	87	87	60
33	['E>=' [74]]	29.6	10.8	73	87	82
34	['E>=' [60]]	26.1	11.9	93	80	58
35	[OR ['H>=' [77]] ['H>=' [74]]]	29.8	10.2	87	100	74
36	[OR ['E>=' [79]] ['E>=' [66]]]	28.2	11.4	60	60	69
Av		**28.3**	**11.8**	**69**	**75**	**74**
SD		***1.9***	***1.0***	***18***	***13***	***14***

Table 4: The equivalent of Table 1 for the best run.

To summarise; the above model succeeds in simultaneously capturing several different aspects of the target behaviour. Of the aspects it does not capture, three stand out. I discuss these below.

- The detailed form of the agents' strategies was often not of a form a human would use in expressing their strategies. They often included expressions that a human would automatically simplify. My model did not include any simplification by agents.
- Unlike the subjects, the agents did not end up by positing and strategies using the O≥*x* predicate and relatively few using the N≥*x* predicate. I postulate that this is due to their early removal as unsuccessful due to the bias against strategies that cost the

maximum amount. I would probably get more realistic results if I allowed the agents to reintroduce these predicates as elaborations upon their strategies at a later state.

- There is a 'tail' of agents with low expected earnings for their final strategy not found in Sonnemans' results, in particular agents 17, 20, 23 and 27 (see Table 4 below). It is noticeable that three of these agents had strategies of the optimal form but had not parameterisation their models effectively. This probably indicates that the subjects performed a more directed and rapid parameterisation of their models than the agents do in my model. This would be consistent with what Sonnemans found in his second experiment in [12].

If the run with the closest match is at all realistic it suggests the following pattern of agent strategy development:

1. Agents start small with simple, uni-predicate strategies and later elaborate them with extra conditions.
2. Agents start by focusing on the form of their strategies and later concentrate upon their parameterisation.
3. Agents judge their strategies against a moderate number of previous games with a strong prohibition against very deep strategies or strategies that do not terminate fairly soon.
4. Agents do a bit of model shuffling, recombining parts of old models.
5. Agents effectively consider quite a large number of hypothetical strategies, but most of these are closely related in form.

7. Discussion

Clearly the model exhibited represents only a step towards an accurate description of the target behaviour observed by Sonnemans. The guidance from cognitive science as to the nature of the learning processes involved is, at present, only suggestive. This results in there being considerable uncertainty about the nature of the learning process, uncertainty that the current observational data does not completely resolve. What the model does do is to point up what information we are lacking about the target behaviour that would enable us to improve the model. In this case one thing that is missing is any direct information about *how* the subjects are developing their models *as* they learn. This could be approached in several ways, including:

- Getting the subjects to specify their strategies using Sonneman's techniques at several stages throughout the learning process, (including before they start);
- The subjects could be asked to specify several strategies under the condition that they would receive the earnings gained by the best performing strategy over a series of runs – this would give a picture of the variety of strategies they were considering;
- Constraining the development of strategies by the subjects to combinations of the operators in the above model using the same 'electronic information board' techniques that Sonnemans uses to capture the strategies, so that some idea about the sort of manipulation and parameterisation of the strategies could be gained;

- Keeping and making available the choice that the subjects made over time so that the dynamics of the learning processes could be checked.

If these were done the process of descriptive modelling could go through another, more accurate iteration. This might highlight more facets that would require experimental or observational investigation. Alternatively it might reveal that some of the inevitable assumptions we had made about the learning process were unwarranted. But in either case we would have learnt something about the actual learning processes involved.

Acknowledgements

Thanks to Joep Sonnemans for answering some questions about his results, to Scott Moss for innumerable discussions, and to Steve Wallis for writing SDML.

SDML has been developed in VisualWorks 2.5.1, the Smalltalk-80 environment produced by ObjectShare. Free distribution of SDML for use in academic research is made possible by the sponsorship of ObjectShare (UK) Ltd. The research reported here was funded by the Economic and Social Research Council of the United Kingdom under contract number R000236179 and by the Faculty of Management and Business, Manchester Metropolitan University.

References

[1] Axtel, R. L. and Epstein, J. M. (1994). Agent-based Modelling: Understanding Our Creations. *The Bulletin of the Sante Fe Institute*, Winter 1994, 28-32.

[2] Edmonds, B. (1998). Complexity and Scientific Modelling. 20th Wittgenstein Symposium, Kirchberg am Weshel, Austria, August 1997.

[3] Edmonds, B. (1999). Modelling Bounded Rationality In Agent-Based Simulations using the Evolution of Mental Models. In Brenner, T. (ed.), *Computational Techniques for Modelling Learning in Economics*, Dordrecht: Kluwer, 305-332.

[4] Hey, J. D. (1987). Still Searching. *Journal of Economic Behavior and Organisation*, 8:137-144.

[5] Holland, J. H.; Holyoak, K. J.; Nisbett, R. E. and Thagard, P. R. (1986). *Induction*. Cambridge, MA: MIT press.

[6] Kogut, C. A. (1990). Consumer search behaviour and sunk costs. *Journal of Economic Behavior and Organisation*, 14:381-392.

[7] Koza, J. R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA: MIT Press.

[8] Montana, D. J. (1995). Strongly Typed Genetic Programming, *Evolutionary Computation*, 3, 199-230.

[9] Moss, S.J. and Edmonds, B. (1998). Modelling Economic Learning as Modelling, *Cybernetics and Systems*, 29:5-37.

[10] Moss, S., Gaylard, H., Wallis, S. and Edmonds, B. (1998). SDML: A Multi-Agent Language for Organizational Modelling. *Computational and Mathematical Organization Theory*, 4:43-69.

[11] Schotter, A. and Braunstein, Y. M. (1981). Economic Search: an experimental study. *Economic Inquiry*, 19:1-25.

[12] Sonnemans, J. (1998). Strategies of Search. *Journal of Economic Behavior and Organisation*, 35:309-332.