

The Dynamic Evolution of Cooperation amongst spatially-distributed agents

Brendan Kitts

*Center for Complex Systems,
Brandeis University, Waltham, MA. 02254.
Ph: (617) 736 2719, Fax: (617) 736 2741
Email: brendy@cs.brandeis.edu,*

Abstract

The aim of this paper is to present a model of how systems can organise into multi-celled communities based around the biological concept of mutualism (Boucher, 1985). The resulting model will be shown to be applicable to understanding phenomena in a variety of domains including economics, sociology, and other environments where simple units are subject to local and global forces. The paper questions the relevance of the Prisoner's Dilemma in understanding the behaviour of these systems, and speculates that swarm-like models may actually comprise an equivalence class. It is suggested that a movement to models more like those traditionally defined in population biology may well provide fruitful insights into the construction of multiple adaptive systems.

Target areas: Parallel models, Collective behaviour, Biological models, Cooperation in Multi-agent systems

Introduction

Life has existed on this planet for a long time; over three and a half billion years. However, for much of our planet's history organisms were nothing like the mammalian forms of today or the striking species of the Cambrian period. Rather, the humble blue-green algae, *dinoflagella*, and other low-level cellular life, have populated our planet for eons. How did these simple organisms, approximately 1.5 billion years ago, begin the heady march towards multi-celled organisation?

The aim of this paper is to present a model of how systems can organise into multi-celled communities, based around the biological concept of mutualism (Boucher, 1985). The resulting model will be shown to be applicable to understanding phenomena in a variety of domains, including economics, sociology, and other environments where simple units are subject to local and global forces. These results suggest that cooperation is a natural limit behaviour of a broad class of adaptive systems, a result which has implications for the way we design and build multi-agent systems.

Model

The model uses a group of cells situated on a spatial lattice, where each cell does nothing more than to take inputs from its local region, and produce outputs to this same region, and by this process we assume it creates energy and keeps itself alive. Unparameterised equations for the system are given below:

$$r_i \dot{(c)} = c_{prod} - c_{cons} + \sum_{n \in \phi(c)} n_{prod} - \sum_{n \in \phi(c)} n_{cons} + f(b(c)) \quad (1)$$

$$e \dot{(c)} = -1 + r_i(c) + \sum_{n \in \phi(c)} r_i(n) \quad (2)$$

$$b(c) = conb_j(c) \bullet \max(|conopt_j(c) - b(cont_j(c))|) \quad (3)$$

$r_i(c)$ is the level of a resource i in a particular cell c , where this level is bounded between 0 and 1, and i can take any value from I to R . $e(c)$ is the energy of a cell and is bounded between 0 and 1. $\phi(c)$ is the set of cells in the neighbourhood of c , with the neighbourhood size $card(\phi(c)) = N$. Each cell consumes and produces a resource defined by $cons(c)$ and $prod(c)$. In equation (1) c_{prod} , n_{prod} etc are meant to be read as "all cells c which produce i ", formally they are defined as: $c_{prod} = e(c)$ if $prod(c)=i$ and is otherwise 0; $c_{cons} = e(c)$ if $cons(c)=i$ and is otherwise 0; $n_{prod} = e(n)$ if $prod(n)=i$, and is otherwise 0; $n_{cons} = e(n)$ if $cons(n)=i$, and is otherwise 0. $f(.)$ is a function which gives a particular "external" payoff value based partly on the behaviour of the cell $b(c)$, but also on whatever other capricious unknown variables that function may happen to be based on. $f(.)$ is meant to capture the idea of "external factors in the environment". In adapting to these forces, each cell has a set of classifier-rule controls, conditional upon the behaviours of their neighbours. $conb_j(c)$ is the response part of the rule (gives the value of the behaviour which the cell will adopt), $cont_j(c)$ is the target cell which c is sensitive to, $conopt_j(c)$ is the condition part of the rule. The j th rule is fired if

the the value for $conoptj(c)$ is the closest to its neighbour's behaviour $b(contj(c))$ (3).

The change in level of a resource (equation 1) should be read as being equal to the energy of all neighbours who are producing i minus all neighbours who are consuming i ; plus the cell's own production of this resource minus the cell's consumption of this resource.

The change in energy of a cell (2) is given by the amount of input resource which is present in the cell's immediate neighbourhood (itself c plus its neighbours n), minus a small constant which is the energy conversion cost.

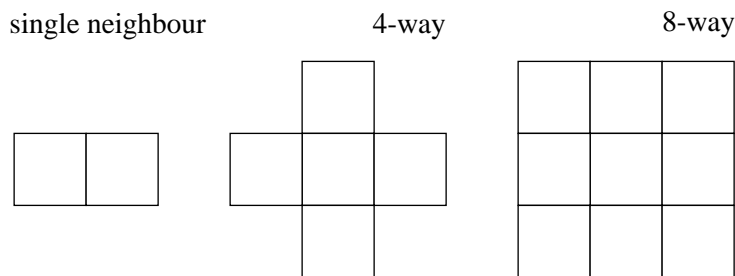


Figure 1: Interaction neighbourhoods corresponding to $N=1$, $N=4$ and $N=8$. Cells consume inputs and produce outputs, changing the quantity of those elements within their own cell. Their production and consumption also affects the nearby cells in their interaction region, such that any consumption in their cell will also deplete the same resource in their neighbour's cell. The arrangement means that one cell may produce output which another cell uses as input, leading to various resource interdependence relationships.

Equation (3) adds a suite of classifier system-style production rules for sensing and responding to the the behaviour of each cells' neighbours. By developing local rules for changing each of their behaviours, cells have the ability to collectively act to increase or decrease the payoff they receive from the function which appears in equation (1), $f(\cdot)$.

$f(\cdot)$ is meant to capture an abstract phenomena, which we shall illustrate through the help of an example, the common sea sponge, *porifera*. *Porifera* are cellular grade animals, which means that they are just a group of specialised cells, with more macroscopic divisions such as organs not present. They are usually immobile, and survive by orientating themselves to the current and filtering nutrients in the water. An interesting feature of these sponges is that their multiple cellular parts have evolved a global structure capable of together enhancing reward: a *funnel* through which water and nutrients can be directed. This idea of individual units selecting various strategies which collectively give rise to reward is captured in the idea of having individual behaviours and a global reward $f(\cdot)$ which is a function of these behaviours.

The particular values for $prod(c)$, $cons(c)$, and the control suite, $con(c)$ are all considered to be changeable parts of the model (Ashby, 1962) and may be subjected to an "evolutionary" process (ie. optimisation) to increase the energy cells each receive from the environment. The optimisation strategy used in the simulations reported here was to have high probability of a cell being replaced

if it had low levels of energy:

$$P_{replacement}(c) = \frac{1}{e(c)} \quad (4)$$

“replacement” meaning that all cell values, including importantly $cons(c)$ and $prod(c)$, are randomised to new values. Other methods such as Stochastic Genetic Algorithms and Real-coded Mutation / Hill Climbing have also been experimented with, to similar effect.

The model was implemented on the Maspar, a 4092 processor parallel mesh computer with a physical architecture particularly well-disposed to implementing the spatial computations required by the model. The next section will discuss some of the results achieved using various parameter values.

Modelling Systems

The basic model outlined above is similar to models in other disciplines such as economics and physics. By selecting various parameter values, many of these models may be simulated.

Ecological Model ($R=20$, $N=8$)

The system can be made to model a hypothetical population of cells under various energy and neighbourhood conditions. Here cells are thought of as the individual units of the model, and resources are chemical structures which exist in the environment and provide nutrients for the cells. “Death and replacement” are implemented using the replacement scheme (4), with global reward $f(.)$ set to zero. The resulting model is designed to show the behaviour of the system with local interactions only under “evolutionary factors”.

Figure 1 shows an example run of the model using these parameters and selection rule. Importantly, the result of these tests show that the number of cooperators (those in the population who exchange resources to each other’s benefit) increase over time, with the number of defectors remaining constant or decreasing. These results will be discussed in detail later in the paper, where the limit behaviour of the system will be analysed explicitly.

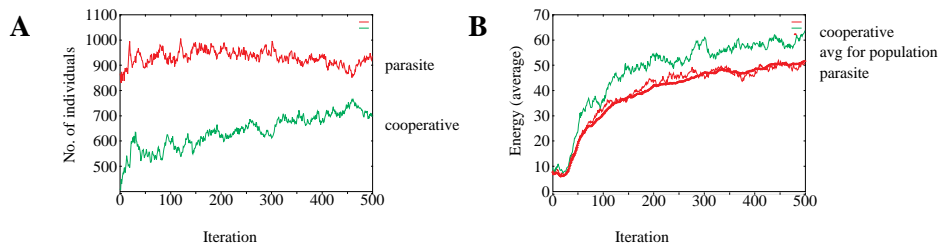


Figure 3: (B) After 1000 trials average energy of parasitic cells deviates slightly around the population mean, where-as the energy of cooperative cells is higher. (A) The number of cooperative cells (due to mutation and replacement) has also

increased, while the number of parasitic cells has remained the same or decreased.

Ising Spin Model ($R=2, N=1..8$)

With some interpretation, the formal processes in the model can be utilized to realise an Ising spin model (Hertz, Krogh and Palmer, 1989).

The Ising model uses the concept of atoms arranged on a lattice, representing the structure of the material. Each atom has a spin and can take one of two values, +1 or -1. The spin of each atom i is determined by summing the interaction strengths in its local neighbourhood $w_{ij}S_j$ (meaning the interactions are correlated with the spins), and adding to this a term representing an external magnetic force applied to the system. The equation for the force on an atom i , is then:

$$h_i = \sum_j w_{ij}S_j + h^{ext}$$

with the spin being determined as:

$$S_i = \text{sign}(h_i)$$

Given uniform exchange interaction strengths, the local field h_i is therefore maximised as:

$$\begin{aligned} h_i &= \max, \text{ if } \forall S_j S_j > 0 \text{ or } \forall S_j S_j < 0 \\ h_i &= \min, \text{ if the positive } S_j \text{ exactly balance the negative } S_j \\ \text{and } h_i &\text{ will have a value inbetween min and max otherwise.} \end{aligned}$$

The characteristics of the model in this paper are similar to the Ising model. Assuming $N=1$, the local exchange strength $e(c)$ is dependent upon the binary “spins” (or binary configuration) of the local cell $cons(c)prod(c)$, and its neighbour $cons(n)prod(n)$, where:

$$\begin{aligned} e(c) &= \max, \text{ if } spin(c)spin(n) \text{ are } 0110 \text{ or } 1001 \\ e(c) &= \min, \text{ if } spin(c)spin(n) \text{ are } 0101, 1010, 0011, 1111, 0011, 0000 \\ e(c) &= \text{value } \min \text{ and } < \max, \text{ if } spin(c)spin(n) \text{ is another combination.} \end{aligned}$$

External force h_{ext} is analogous to the global reward function $f(.)$. Spins are flipped probabilistically according (3), with the resulting model stabilizing into spins of either maximum binary value, although with a slight inconsistency being that the model exhibits only positive interaction forces.

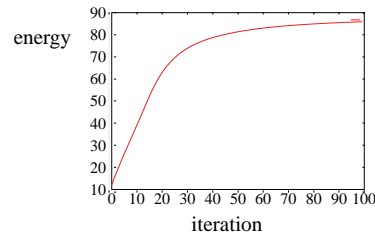


Figure 4: Energy maximisation in the Ising spin analogue.

Macroeconomic Models

In economics cooperation can be seen routinely in the formation of conglomerates and vertical integration practices in Japanese firms. Economic models of such processes have been developed using Cellular Automata (eg. Kirman, 1994, Darley, forthcoming), and there a great deal of scope for emulating these models using the current framework.

Human groups

The model has also been compared against human data from experiments testing cooperation between players in a multi-user game. The main of this data still needs to be verified with a larger sample size, but broadly shows - and is supported by previous sociological studies (eg. Johnson and Johnson, 1989) - that human beings are more likely to cooperate under the following conditions:

- * when individually reward is hard to achieve,
- * when resources are not scarce, and
- * when communication is possible.

These conditions can be artificially induced in the model, in the case of (i) by decreasing payoff cells get from the environment (decreasing $f(.)$), in (ii) by preventing the creation of a zero sum game by allowing replenishment of resources, and (iii) by allowing communication (as opposed to causing cell control target behaviour to be uncorrelated with its actual value, as seen by a cell). In each of these cases, the model exhibits the same qualitative behaviour as in the human experiments - cooperation is either enhanced or impaired.

Further Experiments: Travelling Salesman Problem

Because the particular global payoff function may be provided by a user, the model can be used to solve various practical computing problems where parallelism may be an advantage. One such example is the Travelling Salesman Problem (TSP).

The idea for solving the TSP is to encode cities as behaviours, and restrict the neighbourhood to $N=1$ so that each "city" (cell) is only connected to one other city. The global reward function is then set up to reward short paths and penalise illegal conditions such as duplicate cities. The opti-

misation process encourages cells to form into orderly strands of legal city paths. Initial experiments with the Travelling Salesman Problem have not yet met with complete success, and there is an issue with the coarseness of the optimisation algorithm (ie. it doesn't assign credit well). However, more simple optimisation problems have been shown to be tractable (such as the "stripes" problem), and other ideas on the TSP are still being tried.

The Evolution of Cooperation

The evolution of cooperation has emerged as a somewhat puzzling finding in the literature of iterated games, and particularly the well-known Prisoner's Dilemma model (Axelrod, 1981). The Prisoner's Dilemma is played according to a payoff table like that shown in figure 6A. Each partner elects to either cooperate or defect, and receives a corresponding payoff according to the strategies chosen by both players. In this game defection (exploitation of a partner without giving mutual benefit) can potentially offer greater rewards than cooperation and happens to be the game's Nash equilibrium.

The repeated game with humans, however, usually leads to mutual cooperation, (Poundstone, 1992), and so too it is with the model in this paper (cooperators increase while defectors decrease). Why does cooperation occur in the model system, and does this provide any insight into the Prisoner's Dilemma? In our model a pair of cells (c, n) can be initialised as:

Cooperate-Cooperate ($cons(c)=prod(n), prod(c)=cons(n)$): Both cells supply each other.

Cooperate-Defect ($cons(c) \neq prod(n), prod(c)=cons(n)$): Cell c (the "sucker") provides resources to a neighbour, but neighbour n (the "parasite") does not reciprocate.

Defect-Defect ($cons(c)=cons(n)$): Both cells consume the same resources and so are competitors.

Whilst defection in this framework can initially offer higher payoff, analysis shows that all equations involving defection go to zero in the limit. Similar analysis shows the eigenvalues for a cooperative pair of cells are stable and periodic. Intuitively the reason is that defecting cells eventually kill off their hosts (and so themselves), whilst cooperative cells are able to benignly maintain each other.

	PERCENT OF POPULATION WITH VIRULENCE GRADE					
	I	II	IIIA	IIIB	IV	V
GREAT BRITAIN						
1953	100	0	0	0	0	0
1962-1967	3	15.1	48.4	22.7	10.3	0.7
1968-1970	0	0	78	22	0	0
1971-1973	0	3.3	36.7	56.7	3.3	0
1974-1976	1.3	23.3	55	11.8	8.6	0
1977-1980	0	30.4	56.5	8.7	4.3	0
AUSTRALIA						
1950-1951	100	0	0	0	0	0
1958-1959	0	25	29	27	14	3
1963-1964	0	0.3	26	34	31.3	8.3

FRANCE						
1953	100	0	0	0	0	0
1962	11	19.3	34.6	20.8	13.5	0.8
1968	2	4.1	14.4	20.7	58.8	4.3

Figure 5: Mutations in Myxamotosis virus after initial introduction into hosts in the 1950s. Higher grade strains have lower virulence. (from May and Anderson, 1983; original data from Ross, 1982)

This is in fact consistent with what is known about the dynamics of symbiosis in nature. For example, when the Myxomatosis virus was introduced into the rabbit population in Australia to curtail its epidemic numbers in the 1950s, the virus was extremely virulent, causing high morbidity. But soon after its introduction, less virulent strains of the myxoma virus appeared and began to spread through the population. The drift of myxoma towards various non-lethal strains is shown in figure 5. Thus, since defectors need cooperators to exploit, successful parasites tend to evolve strategies for supporting “cooperators” (May and Anderson, 1983).

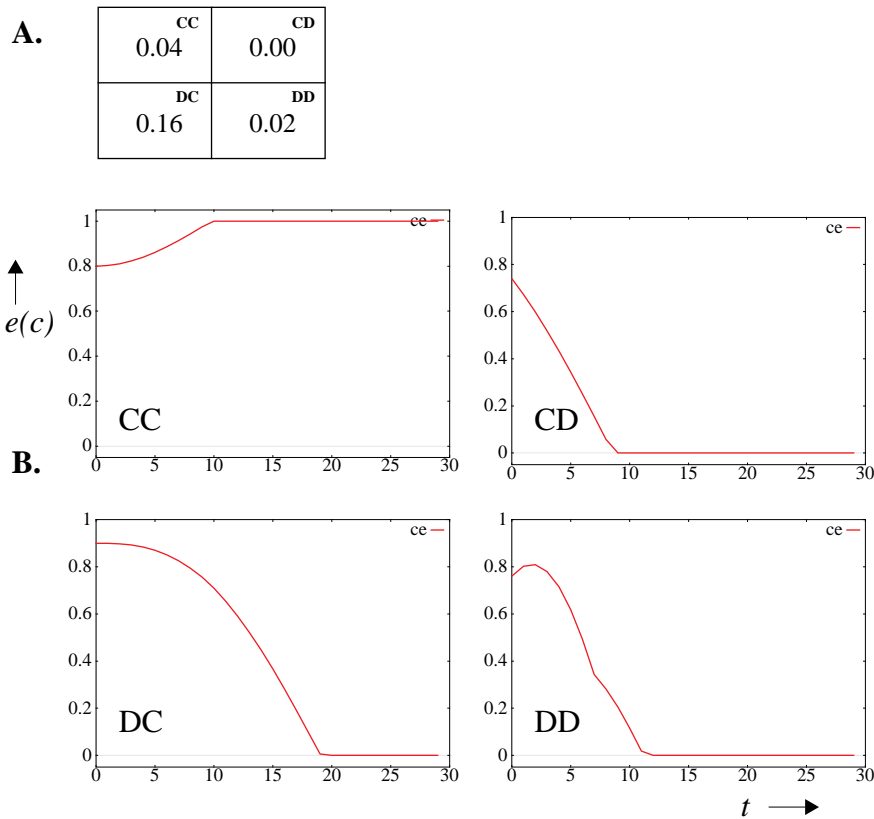


Figure 6: Graphs showing the energy of a cell implementing four particular strategies given initial values $e(c)=0.8$, $e(n)=1.0$, $r1(c)=0.4$, $r0(c)=0.0$, $r1(n)=0.4$, $r0(n)=0.6$. **A:** The initial payoff matrix for this system satisfies the requirements for a Prisoner’s Dilemma problem (Axelrod, 1981). **B:** Iterating the system according to different strategies leads to cell energy going to zero, except in the

case of mutual cooperation, in which energy oscillates through positive values, (in this case CC goes to its ceiling and remains at this value).

Adaptive systems can therefore differ in significant ways from the Prisoner's Dilemma. Payoff matrices typically evolve in time, and choice of strategy interacts with payoff values, such that different payoff variables become dependent upon each other. Indeed, only one half of one percent of the states in the simple linear system analysed here were actually "Prisoner's Dilemma conditions" in the sense defined by Axelrod (1981). Thus, whilst provocative, the Prisoner's Dilemma may not be a good model to study interaction dynamics in adaptive systems.

It seems that dynamical models incorporating the concepts of adaptation, time and interaction effects are needed to effectively capture the richness of multiple, interacting, adaptive systems. It may be more productive, therefore, for designers of multi-agent systems to look to the long and distinguished tradition of models in biology (eg. Boucher, 1984) which have traditionally incorporated these effects, for understanding how multi-agent systems can be synthesized.

Towards an engineering praxis of Massively Distributed Agent Systems: Equivalence classes?

The survey presented above indicates that the model of a system adapting under local and global forces - a model inspired from the biological concept of interacting cells - may represent a broad class of dynamical model. Similar systems have previously been analysed by Chris Langton (the "Swarm Project") including neutron scattering in a nuclear fission reaction and stock market models. A variety of physical systems have also been experimented with, including multiple flocking robot systems (Mataric, 1994; Ünsal and Bay, 1994). This paper suggests that these systems may all share similar dynamics, and so it may be possible to understand the behaviour of them qualitatively.

The practical benefits of this undertaking are significant. Massively parallel multi-agent systems are arguably the future for AI, providing speed, robustness, and scaling properties not deliverable in single processor, centralised architectures. (note that perhaps biology already found this out 1.5 billion years ago). However, in moving to these architectures, engineers are faced with the massive (and computationally intractable) problem of control. How can distributed biological creatures coordinate their various multiple components to achieve computational goals? Presently, researchers at Santa Fe are exploring this issue (eg. Mitchell, Crutchfield and Hrabar, 1994). Treating locally-interacting "aggregate systems" as an equivalence class opens the possibility of broadly understanding the dynamics of these systems in general, and engineering them to exhibit appropriate adaptive properties.

In Artificial Intelligence attempts to create distributed internet agents have been plagued by problems of preventing "defection", (eg. where one agent decides to greedily "cheat" to get the job done). Strategies have ranged from centrally imposed standards and protocols to using punitive agents to deter defection (Rosenschein and Zlotkin, 1994). The analysis in this article shows that as long as agents are able to adapt and no new agents are introduced or changes are made to the environment, the system will eventually settle into a state of mutual cooperation. This result provides guidance on how to design individual adaptive agents, given knowledge about the nature of that

environment (eg. stable/volatile, frequent/infrequent interactions), such that they perform so as to maximise some externally defined global goal. These analyses, which may ultimately borrow heavily from work in biology on population interactions, can allow us to understand what level of cooperation/defection will naturally arise in given environments, and to what extent centralised protocols, with their associated costs on parallelism, should be imposed.

References

Ashby, W. (1962), "Principles of the Self-organising systems", from Von Foerster, H. and Zopf, G., (eds), *Principles of Self-Organisation*, Pergamon Press, NY.

Axelrod, R. (1984), *The Evolution of Cooperation*, Basic Books

Boucher, D. (1985), *The Biology of Mutualism*, Biling and Sons, Worchester.

Brafman, R. and Tennenholtz, M. (unpublished), "On the partial Control of Multi-Agent Systems", manuscript.

Garfinkel, A. (1987), "The Slime Mold Dictyostelium as a Model of Self-Organisation in Social Systems", from Yakes, F. (ed), *Self-Organising Systems: The Emergence of Order*, Plenum Press, NY.

Godfray, H. (1992), "The Evolution of Forgiveness", *Nature*, Vol. 355, pp. 206-207.

Goody, E. (1991), "The learning of prosocial behaviour in small-scale egalitarian societies: an anthropological view", from Hinde, R. and Groebel, J. (eds), *Cooperation and Prosocial Behaviour*, Cambridge University Press, Sydney.

Kirman, A. (1994), "Economies with interacting agents", SFI Working Paper 94-05-030, Santa Fe Institute.

Langton, C. (1996), "The Swarm Simulation System and Individual Based-Modelling", Santa Fe Institute Working paper.

Mataric, M. (1994), "Designing and understanding adaptive group behaviour", Technical Report, Department of Computer Science, Brandeis University.

May, R. (1987), "More Evolution of Cooperation", *Nature*, Vol. 327, pp. 15-17.

May, R. and Anderson, R. (1983), "Parasite-Host Coevolution", Futuyma, D. and Slatkin, M. (eds), *Coevolution*, Sinauer Associates, Sunderland, MA.

Mitchell, M., Crutchfield, J., and Hraber, P. (1994), "Evolving Cellular Automata to perform Computations", *Physica D*, Vol. 75, pp. 361-391.

Poundstone, W. (1992), *Prisoner's Dilemma*, Doubleday, Sydney.

Rosenschein, J. and Zlotkin, G. (1994), "Designing Conventions for Automated Negotiation", *AI Magazine*, Fall, pp. 29-46.

Ünsal, C. and Bay, J. (1994), "Spatial Self-Organization in Large Populations of Mobile Robots," *IEEE International Symposium on Intelligent Control*, pp. 249-254, August 1994.

APPENDIX B

Qualitative analysis of equations

(would like to get the eigenvectors, maybe next time)

APPENDIX C

Data from human subjects

(still being analysed)

GLOSSARY

Symbols used in this paper

$rnd(a..b)$	returns a uniformly distributed random number between a and b
M	maximum level of a resource, c_{mx} (minimum is 0)
c_i	resource which is consumed by a cell (input resource)
c_o	resource which is produced by a cell (output resource)
c_{mx}	level of resource x in the cell's immediate environment
$N(c)$	neighbourhood of c
S	Cell c may consume i , but may also consume values ranging from c_i+S to c_i-S . S is a parameter expressing the "specificity" of resources, with $S = 0$ meaning that resources are completely discrete from one another. All input/outputs are gradually decreased according to how far away they are from the target value, the form of decrease given by function r (response profile).
h_i	force on a particular atom i
E	maximum level of a cell's energy c_e (minimum is 0)
c_b	the behaviour of cell c .
$ccon_b$	behaviour which cell adopts if control conditions are met.
$ccon_{opt}$	the behaviour which cell t should optimally have for the control, con to be triggered.
$ccon_t$	the stimulus for the behaviour change control is cell t 's behaviour, where t is a neighbour of c .
$ccon$	a stimulus-response control which senses a neighbour's behaviour $ccon_b$ and then update's the cell's behaviour, $ccon_b$

$r(offset, range)$	function which returns a value between 0 and 1 to implement a smooth degradation of signal match. r may be any monotonically decreasing function, for instance gaussian or linear. For the purposes of the model a simple linear function was implemented. Note that gaussian or S-shaped curves may have some biological realisability.
$f(.)$	global payoff function. Increases or decreases nutrient levels respectively depending upon the behaviours of the individual cells and whether this meets some “global” criterion.
c	a cell
w_{ij}	exchange interaction strength - local attraction between two atoms
S_i	spin of a particular atom, -1 or +1

NOTES

The first four equations of the model describe changes in input and output volume in the cell’s immediate vicinity, with 3 and 4 practically identical to the first two, but describing corollary changes in the physical neighbourhood $N(.)$. Cells consume from and produce outputs into this neighbourhood at a rate given by the leakage constant L . (figure 1). The energy of a cell c_e is proportional to the amount of input resource in that cell’s immediate vicinity. Energy is subtracted by a small constant which can be thought of as representing energy conversion cost, and if only low amounts of nutrient are present, (eg. < 0.05%), then the cell begins to lose energy. (5)

Each cell may also express a finite number of behaviours c_b (6), and the impact of these behaviours on the environment is given by function $f(.)$. The behaviour which a cell exhibits at any time is the result of an internal control structure c_{con} in which the cell senses its immediate neighbour $c_{con_{tb}}$ and if its neighbour’s behaviour matches a predetermined value $c_{con_{opt}}$, executes a behavioural response to

change its behaviour to that specified in $ccon_b$.

To get some idea of what behaviours are meant to capture, consider the *porifera* the common sea sponge. Sponges are cellular grade animals, which means that they are just a group of specialised cells, with more macroscopic divisions such as organs not present. They are usually immobile, and survive by orientating themselves to the current and filtering nutrients in the water. What is interesting about the sponge, is that its multiple cellular parts have evolved a global structure capable of together enhancing reward: a *funnel* through which water and nutrients may be directed. This idea of individual units selecting various strategies which collectively give rise to a reward is captured in the idea of having individual behaviours and a global reward $f(\cdot)$ which is a function of these behaviours.

In the model discussed in this paper, cells may also choose a defection or coopera-

$$\dot{c}_{m_{c_x}} = -c_e r(|x-i|, S) + f(c_b) \quad (1)$$

$$\dot{c}_{m_{c_y}} = c_e r(|y-i|, S) + f(c_b) \quad (2)$$

$$N(\dot{c})_{m_{c_x}} = -LN(c) e r(|x-i|, S) + f(c_b) \quad (3)$$

$$N(\dot{c})_{m_{c_y}} = LN(c) e r(|y-i|, S) + f(c_b) \quad (4)$$

$$\dot{c}_e = \frac{10c_{m_{c_i}}}{M} - 5 \quad (5)$$

$$c_b = c_{con_b} \bullet \max\left(r\left(\left|c_{con_{opt}} - c_{con_b}\right|, c_{con_{ran}}\right)\right) \quad (6)$$

tive strategy by either utilizing a parasitism or reciprocal resource sharing.

The defection strategy of parasitism (use without reciprocation) is rewarding in the short term, but in the long term the parasitic cell dies, without leaving any of its genetic material to follow. (viruses exception).

The cooperative strategy of mutual exchange and support is long-term stable. Hence, evolution of cooperation in the face of PD-type scenario

NB: situation not PD:

	Neighbour	reciprocate	consume	Cell A	recip-
reciprocate	10 8 10	0 -5 10 (cooperate)			
consume	10 10 10	0 0 0 (defect)			

Take PD through time. Add time dimension. Find consume - reciprocate eventually dies because neighbour dies.

Thus, no dilemma - matrix changes with time. ie. because of the changing resources, what may be a PD matrix changes over time & in this system develops a

Nash equilibria uniquely favouring competition.

Hypothesis: many PD-type scenarios can be reduced to this. esp. biological systems - parasites die off (unless they find a way to keep the host, aren't completely effective in destroying the host, etc.)

that is: the solution to the PD is that in real systems the payoff matrix changes through time, giving less or greater reward to the defection strategies, thus affecting which strategy is most favoured.

The PD itself, (in its purest form), even in the iterated version, is itself intractable because the matrix stays the same. (what is the stable strategy - the Nash equilibria??) The Nash equilibria of mutual defection only exists as a stable point given that payoff matrix. On iterated trials.

Luce and Raiffa (1957), "some hold the view that one essential role of government is to declare that the rules of certain social 'games' must be changed whenever it is unherent in the game situation that the players, in pursuing their own ends, will be forced into a socially undesirable position."

Luce, R. and Raiffa, H. (1957), Games and Decisions, John Wiley & Sons, New York.

For instance, cell c may consume i , but may also consume values ranging from c_i+S to c_i-S . S is a parameter expressing the "specificity" of resources, with $S = 0$ meaning that resources are completely discrete from one another. In the model, inputs and outputs to resources either side of the target x are bounded by S , and all input/outputs are gradually decreased according to how far away they are from the target value, the form of decrease given by function r (response profile).

Simply, it is assumed that resources with closer values will be similar to each other, and may also be consumed by a particular cell, although with gradually less benefit for the cell. The degree of smoothness is given by the "specificity" constant, S , where $S=0$ means there is no smoothness in resourcing, and larger values imply more similarity.

Bolo is an game played between multiple human players and computer controlled robots over a network. The object of the game is to control the most resources in the world. Human beings are allowed to team up or break away at their leisure, and so the game has various strategic aspects which makes it a useful scaled down model of for example competition between rival companies.

Using research from the social science and psychological literature, a number of factors were isolated which theoretically seemed to affect interdependence strategies between humans. This list included:

Resource scarcity which decreases cooperation
Non-equal players which can increase or decrease
Hostility of environment which increases

Ease of Communication which decreases
Iterated contact which increases or decreases.
Kin factors

These factors were implemented in the game of Bolo to collect quantitative data on the emergence cooperation. The results show that cooperation was facilitated in the second, third, and fifth parameter, and discouraged on the first and fourth.

The model can account for three of these five parameter effects. Resource scarcity can be implemented as causing the model to become a zero sum game in which each cell's energy is normalised by the entire population, such that one cell's gain results in less energy for the others in the population. Under these conditions, cells tend not to cooperate, in effect killing each other off, and finally the cell with the most energy dying because of lack of support from neighbours.

Lack of communication was modelled by causing cell control strategies to randomly be chosen (rather than selecting the max). This made it difficult for cells to evolve control strategies based on their neighbours, and impossible in an environment in which multiple behaviours were required.

Hostility was modelled as globally reducing the amount of nutrient available to cells. In these conditions, the only cells which could survive were cooperative colonies.

¶¶-Ø¶¶Ø ¶1Ø1 1Ø Ø ¶ ·Ø 1 · —1Ø ¶
Ø Ø -¶-Ø-1-¶ -¶-1 11 · · ¶-Ø1
Ø -· -1-·Ø-1 ¶.¶-Ø · ¶.1-·
·¶ - 1 Ø —¶¶- · -¶¶-1 1 -·
¶ · · ¶- · ¶.¶ ¶Ø1Ø—
..¶1Ø ¶ - · ¶.-¶-¶Ø·Ø
Ø 1 Ø- 1 · · -¶-Ø 1
- · · ¶¶ ¶¶ ¶Ø Ø ·
· · 1-¶ ¶-¶1¶11 · · ¶.
· · Ø ¶Ø Ø¶¶Ø¶.-Ø- · · -·
-1·Ø -¶ ¶-¶¶-1· · 1¶ 1-Ø Ø
-¶ -¶- 1 -1... Ø.¶ ¶ 11 · 1¶-¶-¶ ·
¶¶· 1 Ø 11-· ·Ø¶Ø¶-¶ · 1 11Ø ¶¶-
¶-¶¶- 1·1Ø· ¶ ¶.¶1Ø· ·ØØ · 1 -¶ØØ
-¶-1 -1 · ¶Ø¶1¶.-Ø· · ¶

Much new research has been forged into these areas of understanding how to harness natural algorithms. For example Denebough has used the behaviour of certain species of ants heaping together food debries, for a sorting model which in fact realises a highly parallel version of a hierarchical clustering algorithm. Recently, Adleman has shown that by careful encoding of DNA molecules, recombination processes can be exploited to create paths in a Hamiltonian path problem, which can then be sifted to find a solution path. In this paper we examine how to solve a Travelling Salesman problem using a distributed model such as the one here.

Because the global payoff function can effectively be sculpted by the designer, it is possible to use the model as a parallel system to compute practical computing problems. In this experiment, cells were specifically encoded as cities, the city given by which nutrient the cell had preference for. Source and destination cities were encoded such that maximum local affinity was given to the closest cities, with decreasing attraction with distance. The global payoff function was then implemented with terms which decreased nutrients to cells which were duplicate, and which increased nutrients proportional to the c/path length. After allowing the population to settle, legal source-destination city pairs could be read off according to their energy, to form a tour path. Using this arrangement, legal tours could be generated which heuristically minimised the path length.

Other NP-complete path problems have already been approached by Adleman, using an approach broadly known as “molecular computation”, (Adleman). In this approach problems are coded as strands of DNA, with natural physical constraints on how they may recombine and assemble. Solutions are assembled as polymer strands, (in fact, every possible solution is generated), and then sifting techniques are used to mark and then sift out the legal solutions. Using this approach, Adleman has solved a 7 city Hamiltonian Path problem, and speculates that the approach may make it possible to tackle various code breaking problems, such as the US Data Encryption Standard.

Related work

A number of previous Alife-style models have been presented already, notably among these Polyworld, and Tierra. This model differs in that it represents an

attempt to incorporate important models presented in a number of disparate domains, and understand characteristics within a common framework.

FURTHER DIRECTIONS

Spatially mobile cells, - CA models already exist which have modelled dispersal. Dispersal perhaps not a very good model of organisations, which is the aim in this case, as most organisations have a spatially fixed internal cellular structure and rely on communication to effect global behavioural changes.

Replicating cells: no attempt made for cells to replicate and/or invade other regions. Population CA have modelled this phenomenon such as invasion of introduced species in existing ecologies, outbreaks of Crown of Thorn Starfish along the Great Barrier Reef, North-Eastern Australia. (Bradbury et. al., 1990).

It is possible that the importance of these factors is small, and cooperation is adequately captured by a spatially localised population. However, parasitism may be a more powerful strategy in populations where iterated contact is less long-term and cells can move about.

Further simulations: Harnessing distributed systems for Practical Computing problems

Using the basic model given in (i), it should be noted that any kind of global payoff function can be used to encourage the model to take on a particular shape of interest. Thus, the cells, cell interactions and global This global payoff function can in effect be hand-coded to represent some problem of interest,

In evolution a similar effect may have been responsible for the prominence of multi-celled organisms in biology. In these cases it is hypothesised that no dilemma really exists - cells which do not play by the rules die off.

The fact that similar model captures behaviour at various levels of analysis including the particles, should be reassuring, since all these systems are governed by the same physical laws. For example, the same laws which allow particles to cluster together may also allow cells to.

The evolution of cooperation should perhaps be best understood as a natural consequence of a system in which the payoff matrix varies as a function of time and players are not trapped into a static, zero win game. Further, as a process recognisable as clumping in particle systems, clustering in cellular systems, and as cooperation in social systems.

Of course, all of this should be old news.

Analysis of the system at steady state (with no evolution) shows that cooperation is the only stable strategy which does not lead to values going to zero.

Depending on the random initialisation of cell resource consumption and production variables, cells may possibly be initialised into any number of different configurations corresponding broadly to “cooperation” and “defection”. Thus for example, a fully mutualistic, cooperating pair of cells will show the property that the resources produced by one are consumed by the other, and the resources produced by the other are likewise consumed by the first. (or $c_{cons}=n_{prod}$, $c_{prod}=n_{cons}$). A “sucker” however, will be any cell which supports another without receiving any benefit in return. Finally, two cells may both mutually defect by consuming the same resources. The respective payoffs for each of these behaviours vary depending on the state values of the model, and will be discussed below.
for instance, changing from a Prisoner’s Dilemma at $c1=...$ to a different at .

This broad kind of relationships is designed to reflect the situation of symbiosis in nature, where different species have different metabolic, behavioural, and spatial characteristics, and so can affect each other’s resource environment in different ways.

Detailed analysis of the system with $N=1$ shows that assuming a closed system

They may produce resources for a neighbour, but that neighbour may not reciprocate (Sucker, $c_{cons}\neq n_{prod}$, $c_{prod}=n_{cons}$)

They may use resources from a neighbour, but not reciprocate (Parasite, $c_{cons}=n_{prod}$, $c_{prod}\neq n_{cons}$)

Both they and their neighbour may use the same resources and fail to reciprocate (Defector, $c_{cons}=n_{prod}$, $c_{prod}\neq n_{cons}$)

As a consequence of the simplistic model assumptions, Individualists who produce and consume the same resource are also possible in the model, although rare in nature! These cells can generally be ignored and don’t affect any of the major findings.

Further experiments: Travelling Salesman Problem

Because the particular global payoff functions are able to be provided by the user, the model can be used to solve various practical computing problems where parallelism may be an advantage. One such example is the Travelling Salesman Prob-

lem.

The idea for solving the TSP is to encode cells with complimentary input-output relationships as source and destination cities. The neighbourhood is restricted to a size of 1, and a global payoff function implemented which gives reward for short paths. The cells are then allowed to optimise energy. Legal tours fragments can be read off as pairs of high energy cells. Examples of tours generated using the method are provided in figure 5, and further details are given in Appendix B.

APPENDIX A

Travelling Salesman Problem

Encoding

For every cell in the population, each is randomly assigned one of the following initialisations (encodings):

$$c_i = \text{rnd}(0..T) 2T$$
$$c_o = c_i + T$$

or

$$c_i = \text{rnd}(0..T) 2T + 0.5T + \text{rnd}(0..T)$$
$$c_o = c_i - T$$

where T is the number of towns (cities). The first initialisation corresponds to an encoding as a source city, and the second to an encoding as a destination. The initialisation is created so that a source cell as an affinity relation with a destination cell, the actual source-destination coding being determined by the closeness of the cities. This local affinity relation means that close cities will tend to attract each other. Finally, a global reward function is introduced:

$$f(.) = 100/\text{path length}$$

which increases nutrient levels to cells as a function of shortness of global tour length. $f(.)$ may also be used to decrease nutrients to duplicate cities, and so on.

Optimisation

Finally, the cells are subjected to mutation-driven optimisation, in this case the algorithm used was:

$$p(\text{mutation}) = 1/\text{energy}$$

and allowed to run for a number of iterations. Legal tours can then be read off as source-destination pairs with “high” energy. (in this case the arbitrary threshold used was $c_e > 10.0$).

The approach of using distributed cells for computing combinatorial problems has been pioneered by Adleman in a series of remarkable experiments which were largely the inspiration for this work. Adleman encoded DNA molecules as city labels and allowed them to recombine into paths which solves the Hamiltonian path problem. The reader is referred to this author for more fascinating applications of molecular computation.

Behaviours are internal changes which cells can learn (or evolve) which may also affect the environment in ways, analogous to the way the sponge changed the physical flow of fluids, and hence the volume of nutrients passing through it.

Thus for instance, in regions of Africa where trypanosomiasis is endemic, indigenous ruminants suffer mild infections and with low morbidity, while important ruminants suffer virulent infections that are usually fatal if untreated, (Alison, 1982). Finally, Dobson (1983), in a survey of 300 parasite-host associations mostly in invertebrate species, found that there was a general (variational) tendency for parasites that are “older” in evolutionary time to be less virulent.

Approximately 90% of all bacteria studied have been identified as carrying plasmids (parasites incorporated into the cytoplasm), and in some cases bacterial recombination can take place with these plasmids. This is also theorised to be the way that cell nuclei were first incorporated into the cell. Levin et. al. (1983) have claimed that: for plasmids to be maintained in natural populations they must carry genes that (under some circumstances at least) enhance the fitness of their immediate hosts or that of cells carrying the plasmid in the population at large.

Empirical studies have shown these new strains of myxoma show decreased reproductive rate. Too high a virulence kills off the hosts too fast, diminishing their capacity to transmit the infection.

Most medical, parasitology texts assert that successful or well adapted parasitic

species evolve to be harmless to their hosts. Simply, all things being equal it is to the advantage of both host and parasite for the parasite to inflict little damage, The same conditions are found in the present model, and lead to nonbenign defectors eventually killing themselves off. The conclusion which can be drawn is a novel one. It suggests that if defectors have , they will evolve ways to support them; and so cooperation will always be a viable strategy in such a system.

Implications for AI

AI has a mandate for adopting massively distributed architectures. Even though silicon operates 5 orders of magnitude faster than biological wet neurons, biology still retains the advantage because of its order of 8 processors working in parallel. Parallelism also provides benefits such as robustness and key speedups in certain areas where there is an algorithmic disposition for parallelism (eg. vision). When entering the realm of distributed autonomous processing however, we quickly run up against the problem of control. How can we expect individualistic cells not to work in ways which satisfy their immediate payoff? The beginnings of a solution are found in biology - cells will *adapt* towards strategies which do not compromise their game payoffs.

The broad conclusions of this work is to suggest that cooperation is *not* a “paradox”, but rather, in adaptive systems, can be a *natural adaptation* for maximising payoff. A large amount of work to this extent has already been meticulously researched through the mathematical biology literature, and supercedes much of this work. A return to a dynamical systems perspective in this case seems warranted. OUT

but it has only been in the last one and a half billion that multi-celled organisms have emerged and proliferated. The process driving the adoption of these complex heterogenous structures appears to have been symbiosis. Symbiosis is a process in which genetically unrelated organisms enter into a relationship in which each organism provides some benefit to the other organism. For example, certain species of ants have learned to tend and protect aphids, which in return secrete a sugary substance which the ants use for food. The evolution of eukaryotes, the acquisition of organelles, and the evolution of biotic complexity have all depended upon the formation of mutualistic relationships between different simple structures (Price, 1986).

This paper will address two concerns. First a model of symbiosis will be described which is applicable to a variety of domains, and in which mutual cooperation is the

limit behaviour of the system. Second, the discrepancies between this model and the Prisoner's Dilemma will be highlighted, and it will be suggested that the adaptive evolutionary systems do not generally face Prisoner's Dilemma.

OUT

This paper will suggest that analogous forces operating in a variety of classes of systems lead to the emergence of cooperation, and further, these forces can be studied and generalities distilled.

Symbiosis and co-evolution appear to be key processes which have been largely overlooked by the adaptive systems community, and even confused by some poor applications of Prisoner's Dilemma style models of economics or repeated games. This paper will attempt to show that.

Garfinkel, A. (1987), "The Slime Mold Dictyostelium as a Model of Self-Organisation in Social Systems", from Yakes, F. (ed), *Self-Organising Systems: The Emergence of Order*, Plenum Press, NY.

Godfray, H. (1992), "The Evolution of Forgiveness", *Nature*, Vol. 355, pp. 206-207.

Goody, E. (1991), "The learning of prosocial behaviour in small-scale egalitarian societies: an anthropological view", from Hinde, R. and Groebel, J. (eds), *Cooperation and Prosocial Behaviour*, Cambridge University Press, Sydney.

Models incorporating the ideas of adaptation, time, and resource-interaction have already been presented and analysed widely throughout the biological literature; indeed, dynamical modelling of populations and parasite-host interactions provides a long and distinguished history (Boucher, 1985).

In virtue of the brain's 100 billion parallel processors, even relatively slow nerve conduction velocities still outstrip, computationally, our current computing hardware.

Bay, J. (1995), "Design of the "Army-Ant" Cooperative Lifting Robot", IEEE Robotics and Auto-

mation Magazine, Vol. 2, No. 1, March 1995.

By optimising their control set, cells will learn which city to switch to given the selection of city of a neighbour. The cells identify which given a source city (neighbour). cells to exhibit behaviour corresponding to where in the path they exist. The global payoff term, $f(.)$, has not explicitly been used up to this point, however, here it can be shown that this can be a particularly versatile part of the model which can allow the user to shape the model to perform computations. The global payoff is a function of cell behaviours and represents the particular physics of the environment, allowing the cells realization of structures such as sieves, funnels, hydrodynamically efficient shapes, and so forth. For example, if the physics of the environment were such that

$$f(.) = 0.5 - |10 - b(c)| / 20, \text{ if } ID(c) \bmod 2 = 0 \\ = 0.5 - |20 - b(c)| / 20, \text{ if } ID(c) \bmod 2 = 1$$

$$f(.) = 0.5, \text{ if } b(c) = B \text{ and } ID(c) \bmod 2 = 0 \\ f(.) = 0.5, \text{ if } b(c) = B/2 \text{ and } ID(c) \bmod 2 = 1 \\ f(.) = -0.5, \text{ otherwise}$$

then the best solution for the cells would be to evolve classifier rules which have $cont(c) = B$ for $ID(c) \bmod 2 = 0$ and $cont(c) = B/2$ for $ID(c) \bmod 2 = 1$, as consequents, and antecedents that anticipate and match their neighbours' behaviours; in other words, the cells should adopt "striped" behaviours. Even though credit assignment in the replacement rule is particularly coarse (ie. a bad rule - or even a good rule which no longer fires because the neighbour's behaviour changed can typically result in the death and replacement of the entire cell). However, after several thousand iterations, the cells robustly tend to adopt the required "striping" behaviours.

Careful selection of $f(.)$ can also allow practical application of the model to various optimisation problems of interest. One such prototypical problem is the Travelling Salesman Problem, a particularly well-known NP-complete problem.

Harcourt, A. (1991), "Help, cooperation and trust in animals", from Hinde, R. and Groebel, J. (eds), Cooperation and Prosocial Behaviour, Cambridge University Press, Sydney.