

On the advantages of information sharing

Michael Lachmann^{1*}, Guy Sella^{1,2} and Eva Jablonka³

¹*Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA*

²*School of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel*

³*Cohn Institute for the History and Philosophy of Science and Ideas, Tel-Aviv University, Tel-Aviv 69978, Israel*

During the evolution of life, there have been several transitions in which individuals began to cooperate, forming higher levels of organization, and sometimes losing their independent reproductive identity. For example, multicellularity and insect societies evolved independently multiple times. Several factors that confer evolutionary advantages on higher levels of organization have been proposed. In this paper we highlight one additional factor: the sharing of information between individuals. Information sharing is not subject to the intrinsic conservation laws that characterize the sharing of physical resources. A simple model will illustrate how information sharing can result in aggregates in which the individuals both receive more information about their environment and pay less for it. This may have played a role in the evolution of higher levels of organization.

Keywords: information sharing; multicellularity; ensociality; public goods; evolution

1. INTRODUCTION

The importance of sharing information about the environment has been noted for many biological systems (Shapiro & Dworkin 1997; Wilson 1971; Rasa 1985; Losick & Kaiser 1997). For instance, myxobacteria have adopted multicellularity as their strategy for survival. Before aggregation individual bacteria measure the state of their local environment. An interactive signalling process involving 100 000 cells ensues, enabling the cells to determine more reliably whether a nutrient shortage is local or whether their nutrient supply approaches global exhaustion. After determining that their available food supply has been exhausted they aggregate to construct fruiting bodies (Kaiser 1993, 1999). Similarly, an ant colony can respond efficiently to information about its environment acquired by many individual ants (Adler & Gordon 1992; Lachmann & Sella 1995; Theraulaz *et al.* 1998). The colony's response will be expressed in terms of the allocation of labour (Wilson 1971), where the large number of ants enables a concurrent operation that increases reliability against environmental contingencies (Oster & Wilson 1978).

In this paper we use the term 'information' in a very simple way: environmental cues carry information to which organisms may react in order to adapt to their environment. We will say that one individual has more information than another if it uses more identical error-prone measurements than the other. We will say information is shared in an aggregate when individuals use the outcomes of measurements performed by other individuals in the aggregate.

Information sharing differs qualitatively from the sharing of a physical resource. When an organism shares information about the environment with others, there may be no immediate cost imposed, no intrinsic loss. In a group of dwarf mongooses, one individual acts as sentry while others forage. A solitary dwarf mongoose would have to invest most of its time looking for predators, and would not survive (Rasa 1985). Assuming that the

members of a band eventually have the same information as a solitary dwarf mongoose, the cost of this information (in this case the loss of foraging time) will be a fraction (one divided by the number of band members) of the cost for the solitary individual. However, when one organism shares food with another it loses exactly the amount of food it gives away. This kind of conservation is intrinsic to the sharing of a physical and/or energetic resource. This distinction, between a physical resource and information, is analogous to the distinction between private and public goods in the economics literature (Taylor 1995).

2. TWO SIMPLE MODELS

To illustrate the effects of information sharing we study several versions of a very simple model. The dynamics occur in discrete time and share the following features.

- (i) Individuals live in an environment that switches between two states E_1 and E_2 . The dynamics of environmental states are described by a Markov process, with switching probability ν per time-step.
- (ii) Individuals perceive their environment through error-prone measurements. The probability of a measurement yielding the correct state of the environment is $1 - e$, whereas the probability of it yielding the wrong state is e . We call e the error probability, and assume that $e < 0.5$. An individual determines its phenotype based on the last M measurements, where M is a parameter of the model called memory size.
- (iii) At each time-step individuals can assume one of two possible phenotypes Φ_1 or Φ_2 . It is assumed that Φ_1 is the 'correct' phenotype for an environment in state E_1 , and Φ_2 for E_2 .
- (iv) A generation consists of T_g time-steps. An individual's relative fitness is the fraction of time-steps at which it had the 'correct' phenotype during a generation. The time-scales are set so that $T_g \gg T_e \gg 1$, where $T_e = 1/\nu$ is the average time during which the environment remains fixed. This relationship enables an individual to infer the state of its environment, and subsequently fitness becomes a measure of an individual's inference ability.

*Author and address for correspondence: Santa Fe Institute, Santa Fe, NM 87501, USA (dirk@santafe.edu).

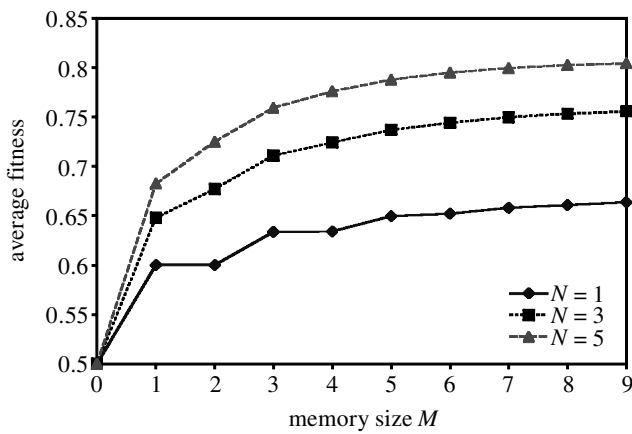


Figure 1. The first advantageous effect of information sharing for individuals in an aggregate: the increase in fitness resulting from having more information. Fitness versus memory size for the optimal strategy in an aggregate, with individuals constrained to one cost-free measurement per time-step. The error probability e is 0.4, and the rate of environmental change ν is 0.05. Each aggregate contains N organisms sharing measurements and memory of M past measurements.

When nothing keeps an individual from performing many measurements (measurements are cheap, quick and independent of one another) then individuals will make enough measurements to cancel the measurement error. This paper, however, will examine cases in which measurement rate by an individual is limited. In the first version of the model, individuals are constrained to one measurement per time-step where a measurement does not entail any fitness cost. In the second version of the model an individual is free to perform many measurements per time-step under the constraint that each entails an additive fitness cost q .

Under these assumptions, what is the optimal strategy for the choice of phenotype? Figure 1 shows a graph of the best fitness possible for different values of M , when the error rate e is 0.4. When only the latest measurement can be used, i.e. $M = 1$, the optimal strategy is to choose Φ_1 when the measurement result was E_1 and choose Φ_2 otherwise. This strategy will yield an average fitness of $(1 - e)$. A method for obtaining the optimal solution for $M > 1$ is described in Appendix A. As M increases the individual can better assess the state of the environment and so its attained fitness increases. For large values of M the fitness asymptotes and does not reach unity, because older measurements become less relevant to the current state of the environment.

We now add an additional assumption:

- (v) Within an aggregate consisting of N individuals, each individual still measures the environment once per time-step, but this measurement together with the rest of the last M measurements are available to everyone.

Figure 1 shows the average fitness of the optimal strategy for a few values of aggregate size N and memory size M , calculated according to the method described in Appendix A. As N is increased, each individual can attain a higher fitness, in some cases greater than the fitness of a single individual with infinite memory. When the number of individuals in the aggregate increases, more recent information becomes available to each, and the fitness

increases. This model illustrates, by way of its construction, the first advantageous effect of information sharing: in an information sharing aggregate each individual has access to more information than it will have in isolation.

Before we examine the second model, we should like to address concerns a reader might have as to the applicability of the first model to biological systems. First consider the memory size. What happens if, as is often the case in nature, memory size entails a fitness cost? In this case fitness will not asymptote with memory size, but instead reach a maximum at a certain size. Increased cost of memory increases the benefit of sharing in an aggregate, because sharing provides an alternative way for the individual to respond to many independent measurements of the environment. A second concern is errors in transmission. Just as there is an error in the measurement of the environment (e), there could be an error in the transmission of the measurement between individuals (e_s). In the analysis in Appendix A we also examine the case with transmission errors. Sharing might not be beneficial when e_s is of the same size or bigger than e . Because signals encode a message, but are not necessarily physically related to it, there is no need to assume that e and e_s are physically related to the message which they encode: signals about an approaching predator do not need to be harder to hear when the predator is harder to see. Sharing will be worthwhile when e is sufficiently bigger than e_s , and we examine those cases.

Now we turn to the second model. Here we consider the effects of information sharing when measurements are costly and the rate of measurement is a variable included in the individual's strategy.

If all individuals in the aggregate behave alike, an individual in an information sharing aggregate that performs n measurements per time-step will pay a fitness cost of qn , and will have information from nN measurements per time-step. The optimal choice of n can be found using the following reasoning. Let $g(x)$ express the fitness gained by an individual having x measurements per time-step, either by measuring or from transmission from others. We assume that $g(x)$ is a monotonically increasing function of x . The fitness of a solitary individual that performs n measurements per time-step is $g(n) - qn$. This fitness will be maximized when

$$\left. \frac{dg}{dx} \right|_{x=n} = q. \quad (1)$$

For N individuals in an aggregate, the fitness for each individual is $g(Nn) - qn$, and will be maximized when (again, we take the derivative of the expression with respect to n and equate it with zero):

$$\left. \frac{dg}{dx} \right|_{x=Nn} = \frac{q}{N}. \quad (2)$$

Figure 2 shows the fitness of a single individual and of aggregates consisting of three and ten individuals, for different values of n , assuming infinite memory. Here we assume infinite memory with no cost, because this is the worst-case assumption for sharing: with increased memory cost the benefits of receiving additional measurements from others increases. One can see in figure 2 that individuals in a sharing aggregate pay less for the same amount of information, because by sharing each individual pays for

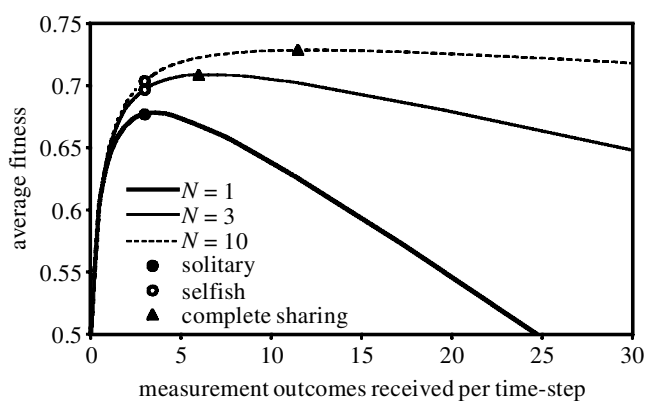


Figure 2. Two advantageous effects of information sharing in the costly measurement model: individuals in an aggregate have more information than in isolation, and pay less for it. Fitness versus measurements per time-step an individual uses. The memory size M is assumed to be infinite. The cost per measurement q is 0.01, and as in the first figure $e = 0.4$ and $\nu = 0.05$. The maximum fitness point for an individual in an aggregate of a given size is shown. The symmetric selfish equilibria are also marked. For selfish individuals the amount of information is equal to that of a solitary individual, but fitness is higher because of reduced cost.

fewer measurements. Each individual can therefore ‘afford’ to gather more information. Under the fitness maximization premise employed above, an individual in an information sharing aggregate experiences two advantageous effects: it has more information than a solitary individual, and pays less for that information. As aggregate size increases, the cost paid decreases.

The fact that information sharing is advantageous to individuals in an aggregate does not imply that it is evolutionarily stable. One problem that may arise is the invasion of selfish individuals that use the information acquired by others but do not contribute to the information pool. Selfish individuals may perform less measurements or abstain from sharing the information they acquire. Consider the stability condition for a group of N selfish individuals, assuming acquired information is non-excludable, i.e. once a measurement has been performed its outcome is available to all (the term ‘non-excludable’ is taken from work in economics; see Taylor (1995)). If $N - 1$ individuals perform n^* measurements per time-step, and the individual under consideration performs n measurements per time-step, this individual’s fitness will be $g((N - 1)n^* + n) - qn$. Taking the derivative with respect to n and equating with zero, we see that it will be maximized if

$$\left. \frac{dg}{dx} \right|_{x = (N-1)n^* + n} = q. \quad (3)$$

Now, consider the symmetrical equilibrium where all individuals are selfish and have identical strategies, i.e. when $n^* = n$. Substituting this into the previous equation gives

$$\left. \frac{dg}{dx} \right|_{x = Nn} = q. \quad (4)$$

In this case each individual has Nn measurements per time-step, and by comparing the above expression with

equation (3) we can see that it has just as much information as the solitary individual, but pays only an N th of the cost. When information is excludable but its sharing entails a cost either through competition between individuals or by way of a physical cost of producing a signal, this model alone does not prevent cheating from invading. The model only exemplifies a benefit that can be gained if cooperation is achieved. A model that will not enable cheaters to invade should have other mechanisms such as kin selection (Hamilton 1964), reciprocity (Axelrod 1984) or population-dynamic effects (Wilson 1987; Eshel 1977; Sella & Lachmann 1999) in addition to information sharing. In Appendix B we present a population-dynamic model in which limited mixing enables information sharing to prevail over selfishness: we assume a constant cost s for sharing a measurement, and aggregates that have a proportion r of genetically identical individuals and a portion $1 - r$ randomly chosen from the rest of the population. We show that if

$$N > \frac{s(1-r)}{qr^2}, \quad (5)$$

then a totally sharing strategy is an evolutionarily stable strategy (ESS). For that strategy the number of measurements is n such that $g'(Nn) = (q + s)/(1 - r + r^2N)$, so that the number of measurements performed per time-step will be higher than in the selfish equilibrium analysed above.

3. DISCUSSION

We have noted in this paper that sharing information is different from sharing other resources. With simple models we showed the interplay between two beneficial effects of sharing information: gaining fitness benefits from having more information, and reducing measurement costs with increasing group size as a result of sharing. We also remarked how the benefits of sharing information by themselves cannot prevent selfish individuals from disrupting sharing in some cases, and that other mechanisms probably come into play in those cases.

During the evolution of life, there have been several transitions in which individuals began to cooperate forming higher levels of organization, and sometimes losing their independent reproductive identity (Bonner 1988; Maynard Smith & Szathmary 1995; Wilson 1971; Buss 1987; Jablonka 1994). For example, multicellularity and insect societies evolved independently multiple times. How can the benefits of information sharing have influenced the evolution of higher levels of organization? One may rightfully claim that an ant colony does not face the same ‘problems’ and is not submitted to the same selection pressures as a solitary insect (Lewontin 1983), and even more so for unicellular organisms as compared to organized multicellular organisms. We suggest information sharing may not only have driven the evolution of higher levels of organization, but may also have played an important role in shaping the changing problems, selection and organization schemes in these higher-level individuals. For example, since the cost of acquiring information decreases with group size, and consequently the amount of information per individual increases as

more information becomes ‘affordable’, the amount of relevant and/or worthwhile information for an individual also increases with the size of the collective to which it belongs. With more information, a collective can react adaptively to a different scale of regularity in the world. Thus quantity of information can become quality: what has been ‘noise’ for an individual can become relevant information for a collective, and consequently the relevant ‘problems’ and applied selection pressures are transformed on the path to a higher level of organization.

We believe that the evolution of collectives and of a higher level of organization involved many different factors. These include the advantages associated with specialization, division of labour and the increased fitness gained by helping kin. Information sharing is another factor, perhaps of no lesser generality or importance. It is likely to have affected the origination of collectives and higher levels of individuality, as well as participated in shaping their evolution. The evolution of a higher level of individuality involved the evolution of mechanisms for acquiring information and sharing it. The simplicity and widespread use of alarm signals indicating the presence of a predator or the depletion of resources, as in the two cases described in § 1, shows that information sharing is already a factor in primitive forms of social organization. To see its effect on higher forms of social organization, one should note the cost–benefit trade-off in such systems: the cost paid per measurement available to the collective decreases with group size, whereas the benefit from making more measurements might level off. This process can lead big collectives to dedicate only a fraction of their individuals to information acquisition—20 sensory cells in a multicellular organism of size 1000 might already give each individual cell 20 times the amount of information available to a solitary cell, for 1/50 of the cost. Thus in advanced forms of social organization evolution of mechanisms for acquiring and sharing information might guide the evolution of differentiation and specialization, as in the evolution of nerve cells and sensory organs in multicellular organisms, or patrolling ants in ant colonies.

We thank C. Bergstrom and ‘Lyons’ in Palo Alto for being indispensable in the development of these ideas, and L. Ancel, D. Ardell, S. Ptak, M. Tanaka, M. W. Feldman, J. Pepper and J. Hughes for many, many comments. This work was supported by NIH grant GM 28016 to M. W. Feldman, and by the Santa Fe Institute.

APPENDIX A

(a) Solving for the best strategy with memory M

(i) *One cell*

In this appendix we give analytical solutions to two related questions:

- (i) What is the optimal strategy, in determining the phenotype in the next time-step, given a sequence of M past measurements?
- (ii) What is the average fitness, across all possible sequences of measurements, associated with that optimal strategy?

The state of the world and the generated measurements are described by the hidden Markov model described in

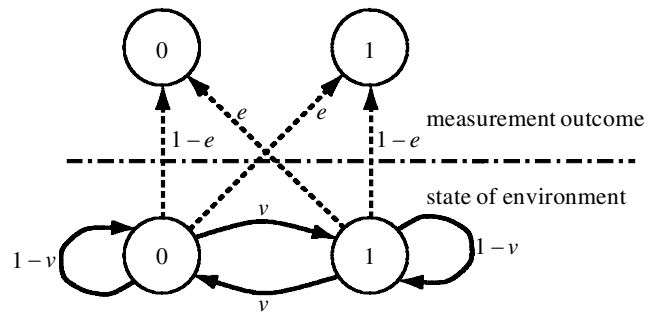


Figure A1. The measurement of the environment described as a hidden Markov model. The rate of measurement errors is e , and the rate of switching of the environment is ν .

figure A1. In order to answer the above questions, we introduce the following notation:

- (i) S_t : the state of the environment at time t , which is either 0 or 1;
- (ii) Π_i , $i = 0,1$: the equilibrium probability of the environment being in state i , which is $\Pi_0 = \Pi_1 = \frac{1}{2}$;
- (iii) a_{ij} , $i, j = 0,1$: the probability per time-step that the environment switches from state i to j , which is $a_{ij} = 1 - \nu$ if $i = j$ and $a_{ij} = \nu$ if $i \neq j$;
- (iv) O_t : the result of the measurement at time t , which in the single cell model is either 0 or 1;
- (v) $b_i(j)$, $i, j = 0,1$: the probability of getting measurement outcome j when the environment is in state i . In the single-cell model this is given by $b_i(j) = 1 - e$ if $i = j$ and $b_i(j) = e$ if $i \neq j$.

Given a sequence of measurements $O_1 \dots O_M$ we would like to calculate

$$\begin{aligned} \alpha_M(0) &\equiv P(S_M = 0 | O_1 \dots O_M), \\ \alpha_M(1) &\equiv P(S_M = 1 | O_1 \dots O_M). \end{aligned} \tag{A1}$$

Then, given $\alpha_M(0)$ and $\alpha_M(1)$, the strategy for the optimal choice of phenotype is

$$\Phi = \text{index}(\max_{i=0,1} \{\alpha_M(i)\}), \tag{A2}$$

and the average fitness associated with this strategy is

$$f_{\text{opt}}(M) = \sum_{\text{all } O_1 \dots O_M} P(O_1 \dots O_M) \max\{\alpha_M(0), \alpha_M(1)\}. \tag{A3}$$

To calculate the average fitness we shall write recursions on $P(O_1 \dots O_M)$ and $\alpha_M(0)$, $\alpha_M(1)$.

First, we derive the initial conditions

$$\begin{aligned} \alpha_1(0) &= P(S_1 = 0 | O_1), \\ &= \frac{P(O_1 | S_1 = 0)P(S_1 = 0)}{P(O_1 | S_1 = 0)P(S_1 = 0) + P(O_1 | S_1 = 1)P(S_1 = 1)}, \\ &= \frac{b_0(O_1)\Pi_0}{b_0(O_1)\Pi_0 + b_1(O_1)\Pi_1} \\ &= \frac{b_0(O_1)}{b_0(O_1) + b_1(O_1)}, \end{aligned} \tag{A4}$$

where we have used Bayes’ law. In the same way

$$\alpha_1(1) = \frac{b_1(O_1)}{b_0(O_1) + b_1(O_1)}. \tag{A5}$$

Finally,

$$P(O_1) = P(O_1|S_1 = 0)P(S_1 = 0) + P(O_1|S_1 = 1)P(S_1 = 1), \quad (\text{A6})$$

$$= b_0(O_1)\Pi_0 + b_1(O_1)\Pi_1, \quad (\text{A7})$$

$$= \frac{1}{2}(b_0(O_1) + b_1(O_1)). \quad (\text{A8})$$

The recursion for $\alpha_t(0)$ can be derived as follows:

$$\begin{aligned} \alpha_{t+1}(0) &= P(S_{t+1} = 0|O_1 \dots O_{t+1}), \\ &= \frac{P(O_{t+1}|S_{t+1} = 0)P(S_{t+1} = 0|O_1 \dots O_t)}{P(O_{t+1}|O_1 \dots O_t)}, \end{aligned} \quad (\text{A9})$$

$$\begin{aligned} &= \{P(O_{t+1}|S_{t+1} = 0) \\ &\quad \times P(S_{t+1} = 0|O_1 \dots O_t)\} \\ &\quad / \{P(O_{t+1}|S_{t+1} = 0) \\ &\quad \times P(S_{t+1} = 0|O_1 \dots O_t) \end{aligned} \quad (\text{A10})$$

$$\begin{aligned} &+ P(O_{t+1}|S_{t+1} = 1) \\ &\quad \times P(S_{t+1} = 1|O_1 \dots O_t)\}, \\ &= \{b_0(O_{t+1})(a_{00}\alpha_t(0) + a_{10}\alpha_t(1))\} / \{b_0(O_{t+1}) \\ &\quad \times (a_{00}\alpha_t(0) + a_{10}\alpha_t(1)) + b_1(O_{t+1})(a_{01}\alpha_t(0) \\ &\quad + a_{11}\alpha_t(1))\}. \end{aligned} \quad (\text{A11})$$

Using a similar derivation for $\alpha_{t+1}(1)$ we obtain

$$\begin{aligned} \alpha_{t+1}(1) &= \{b_1(O_{t+1})(a_{01}\alpha_t(0) + a_{11}\alpha_t(1))\} / \{b_0(O_{t+1}) \\ &\quad \times (a_{00}\alpha_t(0) + a_{10}\alpha_t(1)) + b_1(O_{t+1})(a_{01}\alpha_t(0) \\ &\quad + a_{11}\alpha_t(1))\}. \end{aligned} \quad (\text{A12})$$

Finally,

$$\begin{aligned} P(O_1 \dots O_{t+1}) &= P(O_{t+1}|O_1 \dots O_t)P(O_1 \dots O_t), \\ &= (P(O_{t+1}|S_t = 0)\alpha_t(0) \\ &\quad + P(O_{t+1}|S_t = 1)\alpha_t(1))P(O_1 \dots O_t), \\ &= [(b_0(O_{t+1})a_{00} + b_1(O_{t+1})a_{01})\alpha_t(0) \\ &\quad + (b_0(O_{t+1})a_{10} + b_1(O_{t+1})a_{11})\alpha_t(1)] \\ &\quad \times P(O_1 \dots O_t). \end{aligned} \quad (\text{A13})$$

Using recursions (A9), (A12) and (A13) with initial conditions (A4), (A5) and (A6), $\alpha_M(0)$, $\alpha_M(1)$ and $P(O_1 \dots O_M)$ can be calculated for any of the 2^M possible measurement sequences. By replacing these values in equation (A3) one can calculate the average fitness corresponding to the optimal strategy relying on the last M measurements.

(ii) *A clone of N cells*

Finding the optimal strategy and the fitness associated with it, in the case where information from N measurements is available at each time-step is very similar. The

difference is that now there are 2^N measurement outcomes at each time-step. Considering that, as far as information on the environment is concerned, only the number of cells that obtained each of the outcomes is important and not which cell obtained which result, the number of possible outcomes can be reduced to $N + 1$. We can then denote by j the number of cells that obtained measurement outcome 0, and the number that measured 1 is then $N - j$. Thus, in this case $O_t \in \{0, \dots, N\}$ for every time-step t . Hence, the only modification to the derivations of the last section is in the coefficients $b_i(j)$, where now $j = 0, \dots, N$ and $i = 0, 1$ as before. The new coefficients are

$$\begin{aligned} b_0(j) &= \binom{N}{j} (1-e)^j e^{N-j}, \\ b_1(j) &= \binom{N}{j} e^j (1-e)^{N-j}. \end{aligned} \quad (\text{A14})$$

Substituting these coefficients in the recursions developed in Appendix A § (a)(i), we can obtain both the optimal strategy and its associated fitness, for a clone of size N with memory of M past measurement outcomes.

(iii) *Unreliable signalling*

When signalling is not reliable, the analysis in Appendix A § (ii) must be slightly changed. Now, not all measurements are equally likely to be ‘correct’. In the analysis, we need to distinguish between measurements made by the individual, and measurements gained from sharing. A measurement will now consist of a pair: $O_t = (O_t^I, O_t^S)$, where O_t^I is the own measurement at time t , and O_t^S are the shared measurements. So, now $O_t \in \{0,1\} \times \{0, \dots, N-1\}$. The coefficients in this case are

$$\begin{aligned} b_0(i,j) &= \binom{N-1}{j} (1-e_C)^j e_C^{N-j-1} (1-e)^i e^{1-i}, \\ b_1(i,j) &= \binom{N-1}{j} e_C^j (1-e_C)^{N-j-1} e^i (1-e)^{1-i}. \end{aligned} \quad (\text{A15})$$

Where e_C is the combined error rate of measurement and signalling, which is equal to $e_S(1-e) + (1-e_S)e$. As in Appendix A § (a)(i) it is easy to use these coefficients to obtain the recursions.

APPENDIX B

(a) *Information sharing in populations with limited mixing*

Here we examine the case in which the population is subdivided into separate aggregates, each of which is founded by a single individual. A limited amount of mixing occurs between aggregates: a certain proportion of individuals from all aggregates are put into a ‘mixing pool’, and then redistributed at random back into the aggregates. As a consequence, a fraction of the individuals in a given aggregate will be the offspring of a single parent; the remaining fraction will have entered from other aggregates via the mixing pool and will therefore have different parents. The population is asexual. We will assume that there are two genotypes, which we call cooperators and defectors. Cooperators share their

information, whereas defectors do not—they might even mislead and send false information. Let us use the following notation:

- (i) N : the number of individuals in each aggregate;
- (ii) r : the proportion of individuals in each aggregate that do not get mixed;
- (iii) $f_c(M)$: fitness of cooperators in an aggregate in which there are M cooperators, and $N - M$ defectors;
- (iv) $f_d(M)$: fitness of defectors in an aggregate in which there are M cooperators, and $N - M$ defectors;
- (v) q : the cost of making a measurement;
- (vi) s : the cost of sharing a measurement.

In this model it is fairly easy to calculate invasion criteria for one genotype into a population of another type, and thus we can determine under what cases defection cannot invade a population of sharing cooperators. Determining the steady state of the system in the case of information sharing, when several genotypes can be present at the same time is a non-trivial task, and is beyond the scope of this paper.

When defectors first invade a population of cooperators, i.e. when defectors are rare, the vast majority of aggregates containing defectors will be one of two types: the first are aggregates whose founder was a defector which will have rN defectors and $(1 - r)N$ cooperators; the second are aggregates into which defectors arrived through mixing, and because defectors are rare these aggregates will have one defector and $N - 1$ cooperators. The average fitness of cooperators in a population invaded by rare defectors is $f_c(N)$. Most of the cooperators in such a population will be in aggregates in which there are only cooperators, and only very few will be in aggregates with defectors; therefore the fitness of almost all cooperators will be $f_c(N)$. For defectors to invade, their average fitness has to be bigger than that—even though they might outcompete all the cooperators in their own aggregate. Thus they invade if and only if

$$(1 - r)f_d(N - 1) + rf_d(1 - r)N > f_c(N). \tag{B1}$$

As r increases, the main input to defector fitness comes from the term $rf_d((1 - r)N)$. The function $f_d((1 - r)N)$ decreases as r increases, and is therefore eventually smaller than $f_c(N)$. This is because by supplying wrong information and/or not sharing, defectors eventually hurt themselves in groups in which there are mostly defectors. Therefore, there always exist large enough r , for which defectors cannot invade.

In the following, we examine a specific form of ‘defection behaviour’, the behaviour of sharing only a fraction u of the acquired information. We shall calculate the mixing rate so that limited-sharing individuals cannot invade a population of sharing individuals. Now we introduce the following notation:

- (i) u : fraction of measurements a limited-sharing individual shares with members of its aggregate;
- (ii) n_s and n_1 : number of measurements per time-step done by a sharing and limited-sharing individual respectively;
- (iii) $g(x)$: fitness gain by an individual who receives x measurements per time-step (as defined in this paper).

We assume that g is increasing, and that its derivative g' is decreasing. Using these we can calculate $f_d(M)$ and $f_c(M)$:

$$f_d(M) = g(Mn_s + (N - M - 1)n_1u + n_1) - qn_1 - sn_1u, \tag{B2}$$

$$f_c(M) = g(Mn_s + (N - M)n_1u) - qn_s - sn_s. \tag{B3}$$

Substituting these into equation (B1) and some rearrangement gives the invasion condition

$$(1 - r)g(Nn_s - n_s + n_1) + rg(Nn_s - rN(n_s - un_1)) - qn_1 - sn_1u > g(Nn_s) - qn_s - sn_s. \tag{B4}$$

Let us call a strategy a pair (n_1, u) of choice of measurements per time-step and fraction of measurement shared. We want to see under what conditions a strategy $(n_s, 1)$ will be an ESS. Notice that if $n_1 = n_s$ and $u = 1$, then the left-hand side of equation (B4) is equal to the right-hand side. For n_1 to maximize the left-hand side of equation (B4) at $u = 1$, the following needs to hold:

$$(1 - r)g'(Nn_s - n_s + n_1) + rg'(Nn_s - rN(n_s - n_1))rN - q - s = 0. \tag{B5}$$

If this condition holds for $n_1 = n_s$ then the left-hand side of equation (B4) is maximized for $n_1 = n_s$; so for all other n_1 (with u still equal to unity) the left-hand side of the equation is smaller, and therefore no strategy $(n_1, 1)$, can invade such a population. Solving equation (B5) for $n_1 = n_s$ gives

$$(1 - r)g'(Nn_s) + rg'(Nn_s)rN - q - s = 0, \tag{B6}$$

or

$$g'(Nn_s) = \frac{q + s}{1 - r + r^2N}. \tag{B7}$$

Now we want to find a condition so that no strategy (n_1, u) can invade the population. To do this we first find conditions under which the left-hand side of equation (B4) is a local maximum at $(n_s, 1)$ with respect to (n_1, u) . For this, the derivative with respect to u at this point has to be positive:

$$rg'(Nn_s - rN(n_s - n_1u))rNn_1 - n_1s > 0. \tag{B8}$$

By setting $(n_1, u) = (n_s, 1)$ and using equation (B7) we obtain

$$r^2N \frac{q + s}{1 - r + r^2N} - s > 0, \tag{B9}$$

or

$$\frac{q}{s} > \frac{1 - r}{r^2N}. \tag{B10}$$

Thus $(n_s, 1)$ is a local maximum under these conditions. For it not to be an ESS, another point has to achieve a higher value for the left-hand side of equation (B4), so there has to be another maximum, in this case an internal one. If there is an internal maximum, then the derivative of the left-hand side of equation (B4) with respect to n_1 and u has to be equal to zero. Some algebra gives

$$g'(Nn_s - rN(n_s - un_1)) = \frac{s}{r^2N}, \tag{B11}$$

$$g'(Nn_s - (n_s - n_1)) = \frac{q}{1 - r}.$$

But the right-hand side of equation (B7) is between the values of the right-hand sides of equations (B11). This, together with requirement (B10), means that

$$\frac{q}{1-r} > g'(\mathcal{N}n_s) > \frac{s}{r^2\mathcal{N}}, \quad (\text{B12})$$

or that it has to hold that

$$g'(\mathcal{N}n_s - (n_s - n_1)) > g'(\mathcal{N}n_s) > g'(\mathcal{N}n_s - r\mathcal{N}(n_s - un_1)). \quad (\text{B13})$$

Because g' is a monotonic decreasing function, this is impossible. Thus, a sufficient condition for sharing to be an ESS is equation (B10), or

$$\mathcal{N} > \frac{s(1-r)}{qr^2}. \quad (\text{B14})$$

REFERENCES

- Adler, F. R. & Gordon, D. M. 1992 Information collection and spread by networks of patrolling ants. *Am. Nat.* **140**, 373–400.
- Axelrod, R. 1984 *The evolution of cooperation*. New York: Basic Books.
- Bonner J. T. 1988 *The evolution of complexity*. Princeton University Press.
- Buss L. W. 1987 *The evolution of individuality*. Princeton University Press.
- Eshel, I. 1977 Founder effect and evolution of altruistic traits: ecogenetical approach. *Theor. Popul. Biol.* **11**, 410–424.
- Hamilton, W. D. 1964 The genetical evolution of social behavior, i and ii. *J. Theor. Biol.* **7**, 1–52.
- Jablonka, E. 1994 Inheritance systems and the evolution of new levels of individuality. *J. Theor. Biol.* **170**, 301–309.
- Kaiser, D. 1993 Ronald Thaxter's legacy and the origin of multicellular development. *Genetics* **135**, 249–254.
- Kaiser, D. 1999 Intercellular signaling and the control of cell density and development of *Myxococcus xantus*. In *Microbial signalling and communication* (ed. R. England, G. Hobbs, N. Bainton & D. McL. Roberts). Cambridge University Press. (In the press.)
- Lachmann, M. & Sella, G. 1995 The computationally complete ant colony: global coordination in a system without hierarchy. In *Advances in artificial life. Proceedings of the Third European Conference on Artificial Life*, Granada, Spain, June 4–6, 1995 (ed. F. Moran, A. Moreno, J. J. Merelo & P. Chacon), pp. 784–800. Berlin and New York: Springer.
- Lewontin, R. C. 1983 Gene, organism and environment. In *Evolution from molecules to men* (ed. D. S. Bendall), pp. 273–285. Cambridge University Press.
- Losick, R. & Kaiser, D. 1997 Why and how bacteria communicate. *Sci. Am.* **276**, 68–73.
- Maynard Smith, J. & Szathmàry, E. 1995 *The major transitions in evolution*. Oxford, UK and New York: Freeman.
- Oster, G. F. & Wilson, E. O. 1978 *Caste and ecology in the social insects*. Princeton University Press.
- Rasa, A. 1985 *Mongoose watch*. London: John Murray.
- Sella, G. & Lachmann, M. 1999 On the dynamic persistence of cooperation: how lower fitness induces higher survivability. *J. Theor. Biol.* (Submitted)
- Shapiro, A. & Dworkin, M. (eds) 1997 *Bacteria as multicellular organisms*. Oxford University Press.
- Taylor, J. B. 1995 *Economics*. Boston, MA: Houghton Mifflin.
- Theraulaz, G., Bonabeau, E. & Deneubourg, J. L. 1998 Response threshold reinforcement and division of labour in insect societies. *Proc. R. Soc. Lond.* **B265**, 327–332.
- Wilson, D. S. 1987 Altruism in mendelian populations derived from sibling groups: the hay stack model revisited. *Evolution* **41**, 1059–1070.
- Wilson, E. O. 1971 *The insect societies*. Harvard University Press.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.