Why Information can be Free

Ivana Čače[*][†]

*Cognitive Artificial Intelligence
Department of Philosophy
University of Utrecht
The Netherlands
ivana@phil.uu.nl

Joanna J. Bryson[†]

[†]Artificial models of natural Intelligence (AmonI)
Department of Computer Science
University of Bath
Bath, BA2 7AY, UK
jjb@cs.bath.ac.uk

Abstract

This paper describes a model that demonstrates that sharing knowledge can be adaptive purely for its own sake. This is despite the fact that sharing knowledge costs the speaker in terms of foraging opportunities, and that initially the majority of the population consists of free-riders who listen but do not speak. The population is able to take advantage of the increased carrying capacity of the environment that results from the spread of knowledge, and the free riders are reliably out competed by the speakers.

1 Introduction

1.1 Is Language Costly Signalling?

Many contemporary theories on the emergence of linguistic communication focus on mechanisms that could compensate for the cost of communication to the individual. When transmitted information is viewed as a commodity rather then as a means of influencing the behaviour of the receiver, the evolutionary benefits of linguistic communication do become unclear. The cost for a knowledge transmitter is the cost of giving up a competitive advantage by not keeping valuable information to themselves.

Dessalles (2000) uses agent based modelling to argue against the sharing of information as reciprocal altruism. The starting point of his argument is that the gathering of information is costly and that therefore passing it to others for free is effectively costly too. Dessalles' claims that the results of running his agent model where there is a cost to communicating show that communication cannot be a form of reciprocal altruism under realistic conditions. Only if information were very valuable, the cost for dispensing it very low and if it were easy to detect non-cooperative individuals, would the model favour communication. He then proceeds to explain the benefits of language as a means of forming coalitions and obtaining status within a social group through coalitions.

We will not argue against this last point — Dessalles makes a strong case for how language can serve this purpose. Nor will we argue against any other possible benefits of language that have been brought forward over the years, for example its utility in thinking (e.g. Dennett, 1996).

We do challenge the hypothesis that compulsively sharing valuable knowledge is in itself not adaptive. We have built a model in which agents communicate about food. Agents only have access to food resources they know how to exploit and being told is one way of obtaining this knowledge. The cost of dispersing information is that it enables other, competing agents in the vicinity to gain access to limited resources that otherwise may well have been monopolised by the original knowledgeable, communicating individual. Although there is this individual cost, the population as a whole benefits from the dispersal of knowledge as it effectively increases the carrying capacity of the environment by opening up additional food resources. Our model shows that there are circumstances in which communicating agents outcompete silent, knowledge-hoarding agents, despite the fact that these silent agents hear and exploit all the same knowledge the communicating agents in their vicinity hear and exploit.

In the remainder of this section we give a brief introduction in Agent Based Modelling. In the next section we provide a description of our model, then we report the results of the experiments more finally. We conclude with a discussion and summary of results.

1.2 Agent Based Modelling

Agent Based Modelling (ABM) is used as a research tool in a variety of areas including: social science and economics (Epstein and Axtell, 1996; Cederman

and Rao, 2001), animal behaviour (Hemelrijk, 2000) and complex systems in general (Baray, 1998; Esteva et al., 2001). In ABM, only the agents and their environment are modelled explicitly. The global behaviours being studied in the system emerge from local interactions, both between agents and between agents and their environment, when the model is run. For example, in our models the agents only communicate to their neighbours who are also their immediate competitors for food. This means that for the individual there is an immediate disadvantage in communicating. However, the global effect we show is that over a longer period of time and for the entire population, communication can be beneficial, and will in some circumstances be selected.

Our models were built in NetLogo (Wilensky, 1999). NetLogo is a freely available multi-agent modelling environment, specifically designed for the ABM of natural and social phenomena. The NetLogo world consists of two kinds of programmable agents: an environment divided into patches and mobile agents called turtles. Patches can represent environmental change, such as growing food, while turtles are typically used to represent animals, including humans.

2 Method

2.1 Prerequisites

In this section we describe our model's implementation. We begin with an overview of the general properties of the model such as the relation between the amount of food and the population size. Then we explain how the environment is built and finally we describe agent behaviour.

The model we use has to be suitable for simulating evolutionary pressure and to model the effects of the spread of knowledge through communication. The first requirement entails that the model yields a more or less stable population so that it can be studied over a long period of time. In this context 'stable' means that the population does not become extinct and does not become so small that chance factors can kill off otherwise successful minority species, but also does not explode in size since this would slow the simulation down a good deal.

For the same reason large, rapid fluctuations in population size are undesirable. As a result, the model has to be initiated with values close to the equilibrium. The consequences of events that cause the population to drop dramatically and the ability (and the speed) to get back to a stable population are of

course very interesting, but fall outside the scope of the current research.

In the simulation, time is measured in cycles, one time-step corresponds to one cycle in which all the agents have been activated. So activation is done in a drawing without replacement fashion. Choosing the order in which to activate the agents is resolved by NetLogo itself.

2.2 Basics, the Environment and Food

The number of agents is determined by their birth and death rate, and these in turn are both affected by the replacement rate of the food. The agents reproduce asexually and the reproduction preconditions are chosen such to meet the above-mentioned stability requirement. The offspring function is discussed later in this section, but depends on energy levels. Agents die if they run out of energy or reach maximum age. A maximum age of 50 cycles is imposed to keep a small number of long-lived individuals from influencing the spread of knowledge. In addition, to live the agents need to keep their energy level above zero. They do this by feeding.

Model runs are begun with the values for agent energy and the number of agents close to equilibrium for the given amount of food, including special food. Runs begin with a brief period of population / food oscillation, but these damp quickly to initial equilibrium levels. The population equilibrium then rises slowly as knowledge spreads through the agents, while the food equilibrium stays roughly constant.

The environment consists of 201x201 patch square on a torus space, which is presented on screen as a square. This means agents that walk off any 'edge' of the square will reappear on the opposite edge. On every cycle, energy available in the environment is supplemented in the environment by 'growing food' in a method similar to Wilensky (1998). For every patch, a random number is generated; if it is smaller than the food replacement rate, food is added to the patch.

There are two categories of food: regular food that is available to all agents and six different types of special food with twice as much nutritional value that are only accessible to agents with the corresponding know-how, which enables that food-type's exploitation[1]. A patch can contain only one unit of food. To

---

[1] This idea was derived from one due to Steele (2004), which in turn derived from the Expensive-Tissue Hypothesis (Aiello and Wheeler, 1995). Steele's idea was that language and a larger brain may have co-evolved. Communication enabled agents to exploit richer food sources which in turn allowed them to have larger brains and smaller guts, thus conserving overall metabolic cost.

prevent the special food from clogging up the environment if no agents know how to eat it, special foods can be overgrown by the regular food using the following algorithm.

On each cycle, food growth is accomplished by generating a random number n for all empty patches, and filling those with an n smaller then the replacement rate for the special food with some special food. Then this procedure is repeated for all the empty patches and all patches containing special food using the replacement rate for regular food. This time, if the random number for a patch that is either empty or containing special food is smaller than the replacement rate for regular food, the patch is emptied and refilled with the regular food.

3 Knowledge and Transmission

There are two breeds of agents, agents that communicate and agents that don't. All agents of both breeds understand communication, and will use knowledge received to eat special food if they find it. The only differences between the breeds is that, every time cycle, agents of the communicating breed choose one food type they know about and communicate this knowledge to all their neighbours.

New knowledge comes into the simulated population in a slightly unrealistic way. Agents are all born knowing how to eat the most basic, lower calorie food type. In addition, when the agents are born they have a 5% probability of knowing how to eat any one of the different, special, food types. No learning except from communication is done after the agent is born. Although somewhat unrealistic, this is relatively easy to code. Further it simulates the propensity of young agents to engage in exploration. Imagine if you like that during their first time step of life, agents are more likely to put strange things into their mouths. Again, both breeds are equally likely to acquire new knowledge this way.

The probability of an individual having offspring depends on their energy level. To keep the population size from fluctuating too much we have chosen a conservative offspring function. Even agents with relatively high energy levels do not necessarily reproduce; rather reproduction is probabilistic, though the probability increases with the energy level. When agents with high energy do reproduce, their offspring are at an advantage compared to that of agents with less energy because at birth the parent's energy is split 80:20 with its offspring.

The offspring is always of the same breed as the parent. This may seem too deterministic, but it is easy

to see how a mutation probability biases the ratio between the two breeds. Assume for instance we have a population of 100 individuals of 2 breeds, a 1:10 ratio and a mutation probability of 10% (for explanatory purposes, this is of course too high to be realistic). If the rest of the parameters are such that the population and the ratio between the breeds should stay the same — all individuals hatch once and die immediately afterwards — this is what happens the first time around:

minority breed at $t + 1 = 10 - 0.1 * 10 +$
$0.1 * 90 = 10 - 1 + 9 = 18$
majority breed at $t + 1 = 90 - 0.1 * 90 +$
$0.1 * 10 = 90 - 9 + 1 = 82$

Because we want communication to be the only force influencing evolutionary pressure such a bias would be undesirable.

The agents move across their environment randomly with step lengths distributed according to a Levi-flight distribution. Walking patterns fitting this description have been found in foraging animals as well as in evolutionary optimised foraging agents (van Dartel et al., 2002). The formula describing a Levi-flight distribution is:

$$P(l) = 1/z * 1/l^m$$

Where z is a normalising constant and m is a value between 1 and 3. In our implementation we deviate a little from Levi-flight proper by taking $1/m$ to be 3. This is done to keep the knowledge distribution more localised by keeping step length relatively small. Notice though that the density of the population has no impact on their mobility: any number of agents can be standing in the same patch.

The agents lose a small amount of energy with every time step regardless of the distance travelled. Because agents wander around aimlessly they are presented with feeding opportunities at random, but their chances of actually feeding depend on two things:

1. the know-how they possess, and

2. how many neighbouring agents have the same know-how and may thus be eating the same local resources.

3.1 General Experiment Characteristics

The models are run for a long period of time, typically 12,000 time-steps, corresponding to roughly 350 generations. During this time a number of values are recorded and plotted against the duration of the simulation and recorded separately: the number of agents,

the ratio between the two breeds, the amount of food in the environment, the amount of regular food and the total amount of the special food. The number of things each agent knows, summed up and divided by the total number of agents, is taken as measure of know-how in the population and is also recorded and plotted. While the values above describe properties of the entire population, some properties of individual agents are also recorded. Namely, when an agent dies its ID, age, date of birth, breed, know-how, parent and all offspring are recorded to a file.

Every experiment started out with a population of 10% communicating agents and 90% silent agents. Also, for each experiment, we ran a control version where all conditions were the same except the 'communicating agents' did not actually communicate. They were however still tagged as a different breed from the normal/silent agents.

3.2   Metrics for Evaluation

Recording values as described above provides the means to measure the influence of communication and the opportunity to examine the mechanisms behind it. Values taken from individual agents and related to breed, food distribution and know-how include the following frequencies:

1. Number of offspring.

2. Number of offspring that managed to reproduce.

3. Age at death.

Biological fitness is nothing but reproductive success. Having many offspring is only an indication of fitness if those offspring get to reproduce. For this reason both 1) and 2) are taken as a metric. The age at death is used by Baray (1998) to measure the effectiveness of cooperative behaviour. We use it as an extra metric; it is useful because it correlates with other population properties like offspring survival rate.

There are also two systemic measures: the breed ratios and the environmental carrying capacity. The latter is simply the average number of agents the environment can sustain — Epstein and Axtell (1996) use this to measure the influence of trade. In the models presented here the amount of food in the environment is determined from the start of every experiment, but the amount of food available to the agents depends on the spread of knowledge. With a greater spread of knowledge more food will become accessible, changing the carrying capacity of the environment.

The breed ratios provide a straightforward way of determining how some trait influences fitness. It simply requires checking if individuals possessing that trait take over the population. It is not a sufficient measure because populations can sometimes become extinct despite being well adapted, especially in a model where the total number of individuals lower than about 2000. Because there are fairly significant but essentially random fluctuations of the population, small populations will die out, even if they might ultimately have proved more adaptive. For this reason, if less than 1000 agents are in the initial population, the initial (less than 100) speakers often die out early in the simulation.

4   Description of Runs

The growth rate of the food provides a handle for manipulating population density and consequently the amount of communication. This holds primarily for the growth rate of the regular food, the growth rate of the special food does not have much effect on the population until the appropriate know-how has percolated through the population. Most of the spread of knowledge happens simultaneously with the communicating agents taking over the population; perfect knowledge on the other hand never seems to be achieved.

In order to start the simulation with reasonable values, the amount of special food in the environment has to be approximately the same as the amount of food when the simulation reaches equilibrium. Recall that the equilibrium value depends on the replacement rate and energetic value of the food and on agent life properties. The latter include the offspring function, maximum age and the amount of energy agents lose every cycle. In the first series of experiments the effects of the different food ratios were measured by keeping the growth rate of the regular food fixed and varying the growth rate of the special food. Remember that the 'special food' differs from the regular food by having a twice as high energetic value and a more limited availability.

4.1   Preliminary Results

We ran experiments with different amounts of food in the environment. Under almost all of the conditions the silent agents die out. As expected, with comparatively more special food in the environment the talking agents had more of an advantage. The time at which all the non-communicating agents die out is inversely proportional to the replacement rate of the special food. Only when the amount of special food in the environment is kept very low does the advantage of communicating disappear. This is done by

food replace rate		average population size	
regular	*special*	*with comm.*	*without comm.*
16	1	1367.51	1336.21
16	2	1440.35	1373.81
16	8	2303.71	1447.97
16	16	4267.45	1579.35

Table 1: The average population size after equilibrium is achieved in two different conditions (with and without communication) at four different replacement rates for special food.

setting a (very low) limit to the amount of food, the food only gets replenished if the sum total falls under this threshold.

With communication the carrying capacity of the environment increases with the amount of special food. In the communication-free control conditions there is also a population increase but much smaller. This increase is due to the individuals who are randomly born knowing about special food, even though they can't communicate about it. Table 1 shows the total population size by the end of the simulation (set to 12,000 time cycles, which is usually well after a stable equilibrium is found.) Notice that in the last two cases the silent agents were already extinct at the time the population stabilised enough to get an average, but in the other cases the population stabilised *before* the silent agents died out.
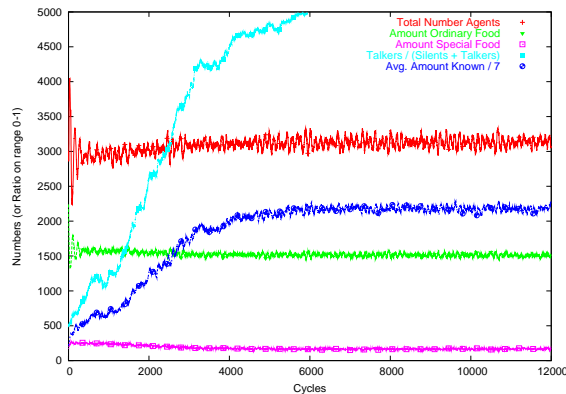

Figure 1: An example run where there are 6 special food types. Points are drawn approximately 10 generations apart.

Analysis shows there is no exclusive or common know-how. That is, for every one of the different things the agents can know, at any point in time, the percentage of agents that have this knowledge will fall in the same interval, yet there is never a point reach where all agents know all things (see for ex-

ample Figure 1). Note that this means only that individual knowledge is distributed evenly at different points in time. Across particular regions of space some know-how may be shared by many agents while some could be held exclusively by a small number of individuals. In fact this is exactly what happens. We have equipped our model with a visual diagnostic tool to enable us to see which agents have the same know-how. Of course, having some type of know-how will lead to a higher energy level only if the agent stumbles upon the corresponding food. The speed of dissemination and the overall availability of knowledge depends on the density of communicating agents. Knowledge spreads faster in an environment that can support more agents. Also the numbers of communicating agents increases faster in an environment that has more special food (food that requires knowledge) in it.

5 Discussion and Future Work

5.1 The Effects of Communication

We have chosen a natural way of modelling the cost and benefit of communication. In our model the costs of communication are individual and temporally local. Also, the costs and benefits (for the receiving agents) are probabilistic. The hearing agent only benefits from the newl knowledge if it stumbles on a patch with the corresponding food. The cost for the communicating agent mirrors this. By compulsively giving out information the communicating agent enables its neighbours to take that same food. There may never actually be any competition if the food is not there. But when there is food in the vicinity, the neighbouring agents will be in the position to take it all. Nevertheless, compulsively giving free information is the evolutionarily successful strategy.

5.2 Selection

It may be argued that what we are showing here is kin selection. The agents live for a short time and they mostly take small steps. Under these circumstances an agent can be expected to have more contact with its offspring or parent than with any other particular random agents. These are simple agents which speak to every neighbouring agent the same, whether they are related or not, or even whether or not they are the same breed. Nevertheless, some biologists argue for kin rather than group selection even where the identity of kin may be 'mistaken', or in this case, not discriminated.

We can see no other explanation for why the silent agents die out except that, since they tend to be near their own relatives, they don't tend to get to know as much, and so they are out-competed for energy. Similarly, although a speaker gives up the advantage of its own knowledge in the short term, because it is likely to be in a community of similar agents, it is more likely to learn knowledge that helps it exploit more different kinds of food. What our simulations show is that, at least in some circumstances, this is the better strategy than free-riding.

There is obviously a great deal of work remaining to be done. It would be interesting to characterise more formally when and why the silents die out and the communicators dominate. We could also count how many of the communications going on are between relatives as a proportion of the whole. It would also be interesting to modify the model to explore things such as what happens if agents *do* restrict their communication to close relatives, practice deception, or just accidently, through their own ignorance, communicate useless or even harmful knowledge.

Nevertheless, the work as it stands brings interesting challenges to some existing Evolution of Language theory. For example, every theory on the emergence of language that presupposes that the cost to communication automatically means communication is not in itself adaptive needs to be reevaluated.

6 Conclusion

We have demonstrated that the propensity for communication can have a selective advantage despite being costly, provided that it has sufficient benefit to the community of speakers. The *act* of communicating may be costly to an agent, but the *propensity* to communicate will benefit the agent if it is consequently likely to learn from its children, parent, siblings and cousins. Further, we have shown that free riders who understand what they hear but do not share their own knowledge not only fail to inhibit the selection of free communicators, but will in the long term be out-competed if they tend to pass on their propensity for not sharing their knowledge and to live near their kin.

As we stated in our introduction, our results in no way challenge whether there are other selective forces that have affected the evolution of language, particularly language as we know it. But we have conclusively shown that arguments in favour of such mechanisms as selection for prestige cannot rely for evidence on free communication being non-adaptive. Communication may still have been the first selective advantage of language.

References

L. C. Aiello and P. Wheeler. The expensive-tissue hypothesis: The brain and the digestive-system in human and primate evolution. *Current Anthropology*, 36(2):199–221, April 1995.

Cristobal Baray. Effects of population size upon emergent group behavior. *Complexity International*, 6, 1998.

Lars-Erik Cederman and M. P. Rao. Exploring the dynamics of the democratic peace. *Journal of Conflict Resolution*, 45(6):818–833, 2001.

Daniel C. Dennett. *Kinds of minds: Towards an understanding of consciousness*. Weidenfeld and Nicolson, 1996.

Jean-Louis Dessalles. Language and hominid politics. In Chris Knight, Michael Studdert-Kennedy, and James Hurford, editors, *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, pages 62–79. Cambridge University Press, Cambridge, UK, 2000.

Joshua M. Epstein and Robert Axtell. *Growing Artificial Societies: Social Science from the Bottom Up*. Brookings Institution Press, MIT Press, Cambridge, MA, 1996.

Marc Esteva, Julian Padget, and Carles Sierra. Formalizing a language for institutions and norms. In Jean-Jules Meyer and Milinde Tambe, editors, *Intelligent Agents VIII*, volume 2333 of *Lecture Notes in Artificial Intelligence*, pages 348–366. Springer Verlag, 2001. ISBN 3-540-43858-0.

Charlotte K. Hemelrijk. Towards the integration of social dominance and spatial structure. *Animal Behaviour*, 59(5):1035–1048, 2000.

James Steele. What can archaeology contribute to solving the puzzle of language evolution?, April 2004. plenary talk at *The Evolution of Language*.

M.F. van Dartel, E. 0. Postma, and H.J. van den Herik. Universal properties of adaptive behaviour. In H. Blockeel and M. Denecker, editors, *Proceedings of the 14th Belgium-Netherlands Conference on Artificial Intelligence (BNAIC '02)*, pages 59–66, Leuven, Belgium, 2002.

Uri Wilensky. The netlogo rabbits/grass/weeds model, 1998. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Uri Wilensky. *NetLogo*. Evanston, IL, USA, 1999.