

Information Theoretic Incentives for Social Interaction

Christoph Salge¹ and Daniel Polani²

¹ University of Hertfordshire, Hatfield AL10 9AB, UK,

`c.salge@herts.ac.uk`

² `d.polani@herts.ac.uk`

Abstract. A first step towards social interaction is to observe other agents and their actions. The concept of “Relevant Information” is used to argue, from an information theoretic perspective, why it would be beneficial to observe other agents, and why *observing* their actions should be relevant to me, even if their actions are not. A simple grid world model illustrates those points, for a simple information gathering task, and shows how to utilise this information to increase an agent’s performance.

1 Introduction

Information Theory [Sha48] has been a powerful tool in advancing our understanding of adaptive behaviour from an agent centred perspective. Modelling the world as a set of random variables, and applying an information based metric, lead to the development of “low-assumption” behaviour strategies, such as “Empowerment” [KPN07] or “Homeokinesis” [ABD⁺07]. Just accessing their own sensors and actuators, agents driven by those strategies try to maximise information and channel capacities, but they struggle to qualify information.

Shannon information is in itself meaningless [Neh99], but it is possible to quantify the relevance of information in regard to the agent’s action. If there is a optimal action strategy (in regard to a fitness function, such as survival), and if this strategy differs depending on the state of the environment, then there is a certain amount of information the agent has to obtain to determine this strategy. In [PMK01] Polani demonstrates that the amount of this information can be measured as the mutual information [CT91] between the agent’s actions and the environment. But since the mutual information tells us how much information one random variable contains about another this means that the “Relevant Information” could also be obtained from the actions of the agent itself.

If we now deal with agents that have a limited capacity to process information, then those agents have to select from what source they obtain their “Relevant Information”. But, if there are other agents, obtaining similar information, and then acting upon it, observing another agent might be a good strategy. Several biological examples for this exist, and Danchin [DGVW04] argues that this might be due to “Inadvertent Social Information”. We wanted to examine this within our formal framework, and also study the information

theoretic properties of agent’s actions, to better understand the incentives to observe others.

2 Non Social Agent Model

To study those phenomena we created a grid world model, as described in Fig.1, where an agent has a simple information acquisition task.

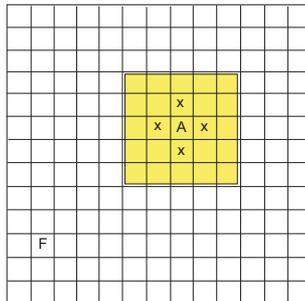


Fig. 1. A periodic gridworld of the size $n \times m$ (in this case 13×13), A indicates the position of the agent, F the position of the foodsource, both are random at the start of the simulation. Its task is to locate the food source F. Every turn the agent can move to a neighbouring cell, indicated yellow, and gets new sensor inputs. The yellow area marks all the cells visible to the sensors of agent A if it has a sensor range of 2. The cells can be perceived as “empty” or as containing “food”. It then decides again where to move, until it locates the food source.

The agents determine their actions by using an internal memory that stores information about the world. The internal memory is an array of values, each associated with a cell in the world. Those values store the probability for a cell to contain food, given the experience of the agent. Initially all cells have the probability of $\frac{1}{n \cdot m}$. However, as the agent moves around, it discovers that some cells are empty or contain food. The distribution of probabilities is adjusted by setting either the probability of a cell to zero in the case that there is no food in it, or to one in case it does. In both cases the other cells are normalised, so the sum of probabilities is still one. The uncertainty of the food source position is reflected by the probability distribution and can be measured in terms of entropy $H(F)$, where F is a random variable encoding the expected position of the food.

$$H(F) = - \sum_f P(F = f) \cdot \log P(F = f)$$

2.1 Infotaxis Search

To generate the agents behaviour, we adapted a method called “Infotaxis” [VVS07], which provides an information theoretic for how a moth could use the

sparse information provided by their olfactory sensors to determine the source of certain smells. The main idea of infotaxis is to act not in a way that gets you closer to you target, but in a way that increases your expected information gain.

To determine which way to go the agent selects the move which has the highest expected reduction in entropy, according to its internal Bayesian model of F , the random variable encoding the food source. It is important to note that the calculation of the expected entropy reduction of F is done by using the current distribution of F . So the expected reduction of entropy is based on the current agent’s “knowledge” about F .

To formalise, we define the set that contains the positions w of all the cells of the grid world as $\mathcal{W} = \{w = (i, j) | 0 < i < n + 1, 0 < j < m + 1\}$. The values i and j are the coordinates of the position on the grid world. Note that $|\mathcal{W}| = n \cdot m$. Also, the random variable F that encodes the food source position, from the perspective of the agent, uses \mathcal{W} as alphabet. Also, since we are considering a world with periodical boundaries both sides of the equation $(i, j) = (i + n, j + m)$ denote the same position. Now, depending on the position of the agent w_A , there is a set \mathcal{S} that includes all the positions that are visible to the sensor of the agent. Seen from the current position there is another set \mathcal{S}_a that is defined as the set of sensor inputs after the agent takes action a from a set of possible actions \mathcal{A} .

To calculate the expected entropy reduction $\Delta H(a)$, depending on the action a , two main cases have to be considered. In the first case the actual location of the food source $f \in \mathcal{W}$ could be in \mathcal{S}_a , the sensor range after the action a was taken by the agent. The agents assumes that this occurs with the probability of $P(F \in \mathcal{S}_a) = \sum_{f \in \mathcal{S}_a} P(F = f)$ in reference to our internal variable F . In that case our current uncertainty would be reduced to zero, so the reduction in entropy would be the full $H(F)$.

The other case assumes that the location f of the food source is not in \mathcal{S}_a . This occurs with a probability of $1 - P(F \in \mathcal{S}_a)$. In that case we have to calculate an updated probability distribution of F , which we will call F_a . We set all $P(F_a = f) := 0$ for all $f \in \mathcal{S}_a$ and then normalise the other probability assumptions by dividing them by the sum of all probabilities.

$$\forall (f \notin \mathcal{S}_a) : P(F_a = f) = \frac{P(F = f)}{\sum_{w \notin \mathcal{S}_a} P(F = w)}$$

The updated version of F_a can then be used to calculate the reduction of entropy $H(F) - H(F_a)$ in the second case. Put together the expected reduction of entropy is:

$$\Delta H(a) = P(F \in \mathcal{S}_a) \cdot H(F) + (1 - P(F \in \mathcal{S}_a)) \cdot (H(F) - H(F_a))$$

So every step the agent selects randomly an action a that maximises $\Delta H(a)$.

2.2 Results of Infotaxis

As a measure of performance we record the time it took the agents to locate the food source. On average, the agents with the infotaxis behaviour outperform

agents that chose their direction at random by a significant factor. For a 25×25 world, the average search time for the location of food, measured over 50000 trials, is around 76 turns for infotaxis agents (see Fig.2.b), and around 450 turns for random walk agents, which we implemented as a baseline comparison.

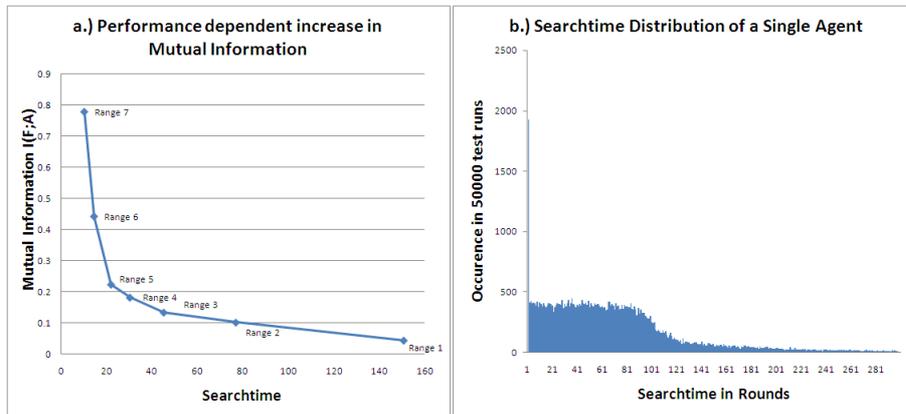


Fig. 2. a.) shows the relation between performance, measured as average searchtime, and the amount of mutual information between the agent’s action and the position of the food source. The different values are created by agents with different sensor ranges b.) shows the distribution of seachtime, the time it takes for the agent to locate the food source. Number roughly approximates the theoretical optimal distribution.

We also record the states of A , the actions of the agents, together with F , the relative location of the food source. This makes it possible to calculate the joint entropy $H(F, A)$, which can in turn be used to calculate the conditional entropy as $H(F|A) = H(F, A) - H(A)$. This allows us to calculate the mutual information as $I(F; A) = H(F) - H(F|A)$, the value that indicates how many bits of information an action of another agent tells us, on average, about the location of the food source. The results for the different sensor ranges can be seen in Fig.2.a. If we compare the information gained from the action of an agent with sensor range two ($\approx 0.1bits$), to the information gained from observing a single cell ($\approx 0.01bits$), than the information from the agent is significantly larger.

This shows that the agent’s action yield relevant information, and considerably more than the observation of a cell. This should give an agent, that wants to obtain similar relevant information and has a limited processing capacity, an incentive to discriminate towards information from other agents.

Furthermore, if we vary the sensor range of the agents in the prior experiment, we can create agents that perform better because they have access to more information in the same time. As Fig.2.a shows, the increase in range, and the

ensuing increase in performance also increases the amount of relevant information encoded by the agent. Therefore, the observed agent also has an incentive to encode as much relevant information as possible, since this is often necessary to obtain a better performance level.

3 Social Agent Model

In the second part of the experiment we want to demonstrate that the relevant information is not only encoded by the agents, but can also be used, with relatively simple mechanism, to increase the performance of an observing agent.

The extended version of the Fishworld model populates the grid world with several agents at the start of the simulation. All of the agents try to find the same food source. Also, in addition to the current two sensor states, corresponding to a cell being empty or containing a food source, the sensors of the agents can now assume new states that indicate that another agent moved into the corresponding cell and which direction it came from. So the four new states are {agent that moved in from the north, ... south, ... east, ... west}. So, if an agent senses another agent within its sensor range, it will be able to perceive the direction it entered the cell from; the agent is able to see the last move of that other agent.

Since we now want to evaluate what an agent could possibly learn from the other agent’s actions we had the agents update their own internal model of the food source F depending on the other agent’s actions. This adjustment of probabilities can be comfortably integrated into our existing infotaxis search. Note that for the described simulation *all* agents are equipped with those new “social” abilities and all of them use the other agents’ actions to update their internal world models.

Note that none of the agents take an active interest in communicating and one agent’s actions does not influence the performance of another agent, so there is no competition for food resources and no collision.

3.1 Bayesian Update

F is our current internal probability assumption of the location of the food source, and $a \in \mathcal{A}$ are the states of the random variable A that encodes the last move action of another agent we are observing. We can then use Bayes’ Theorem to update the probability distribution of F . What we are interested in is the probability of the food source to be in a specific location given the evidence of another agents action and position $P(F = w|A = a)$. According to Bayes’ Theorem this could be calculated for every cell f of the environment as:

$$P(F = f|A = a) = \frac{P(A = a|F = f) \cdot P(F = f)}{P(A = a)} \quad (1)$$

$P(F = f)$, the a priori probability, is the internal memory of the agent mapping the probability distribution of F , as gained by their own experience so far. $P(A = a)$ is the probability that the agent takes the move action a . Since the

grid world in all our example was symmetrical to rotation the probability is close to $1/4$ for every a ; Our statistics also confirms this. The term $P(A = a|F = f)$ is the probability of another agent performing action a if the food is in position f . Note that the position f in this case is calculated in relation to the position of the observed agent. So the question we are asking is for example “If the food is 3 cells north of the agent, what would be the probability of the agent performing move action a ”. We then look at all the cases where in the past an agent has been 3 cells south of a food source, and determine which action it took.

To obtain those statistics for the computer simulation we observed 50000 non-social infotaxis agents searching for the food. Note that the agents we used were non-social and thus “blind” to the actions of other agents. They behaved as described in the “Non Social Agent Model” part of this report. So even though all the agents in the simulation now have the ability to sense other agents and update their internal world models they still calculate the update as if they assumed that all others were non-social agents. After the agent updates F , it resumes the infotaxis behaviour described in the non-social agent model.

For our current simulation all agents were equipped with the new mechanism, but only the performance of one singular agent, chosen at random, was measured. All agents started in their search at the same time and where initialised with no knowledge about the world. Agents that had successfully located the food stopped moving and were no longer perceivable by other agents. This was done, because it would have been very trivial for another agent to infer from seeing another non-moving agent that the food must be within sensor range of that agent. So the agents could not see agents that knew where the food was.

3.2 Results

If we run such a simulation, populated by 10 agents with a sensor range of 2, the average searchtime is around 65 for the social agents, as compared to around 76 turns for the non-social infotaxis strategy. Since we did not include any collision detection, or any other means for the agents to interfere with each other, the searchtime for the non-social agents remains unchanged in regard to the amount of other agents in the simulation.

In the social agent simulation a further increase of the number of agents reduces the average searchtime further, down to a minimum of 48.9 turns for 95 social agents, as seen in Fig. 3. An increase in the number of agents then leads to a slow rise of the average searchtime again. A simple calculation shows, that the optimal searchtime for a non-social agent is 57.6 turns. So, the performance is not only increased by observing others, but it is increased beyond the level of any non-social strategy. Therefore, the observed agent must yield information not obtainable otherwise.

4 Discussion

The mutual information between the agent’s actions and the environment, which is used to calculate the amount of “Relevant Information”, can be interpreted

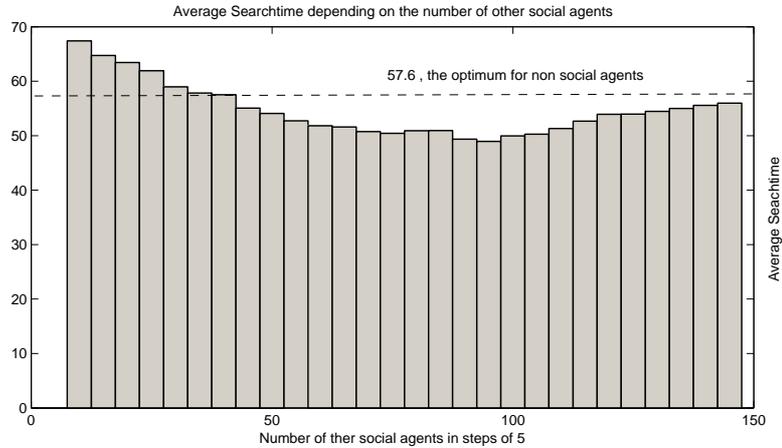


Fig. 3. The average searchtime over 5000 trials in a social Fishworld simulation of the size 25×25 . Measurements were taken for several populations sizes in steps of 5

in three ways: It is the amount of information the agent needs to obtain to determine its optimal strategy, it is the amount of information it needs to encoded in its actions to act upon the obtained information, and it is also the amount of information I can gain about the environment from looking at the agent. Therefore, it is no surprise that the agent’s actions contain relevant information. But because the agent has to encode the amount of information into its actions (which in general have a much smaller state space than the environment) the agent is faced with an information bottleneck [TPB99]. This forces the agent to “digest” the information, and leads to a higher density of observable relevant information, in comparison with other parts of the environment, such as empty cells.

Also note that it is not only possible to determine the amount of relevant information for an optimal strategy, but also for any suboptimal performance level [PNMK06]. Increasing an agents performance level, as seen in Fig.2.a, often necessitates an increase in encoded information. So it is in the agent’s own interest to increase the amount of encoded relevant information as a side product of increasing its performance. Interestingly, this illustrates why it is more beneficial to observe an agent that performs better.

All those effects so far aRE results of the agents inner motivation to increase its information gain and its own performance. So far, the agents have no active interest in communication, or any specialised mechanism to even identify another agent. Also, the actions of one agent could not influence the performance of another, since there were not scarce resources or a possibility for collision. But as the social model shows, the information gained from observing others could lead to a significant increase in performance, surpassing any non-social strategy. And since the agents could not obtain information from others and not act upon

it (if they wanted the benefit of that information), the agents had to react to the actions of others, leading to a rudimentary form of social interaction.

If we understand the ability to identify other agents, and possibly their actions, as a prerequisite for other, higher levels of social interaction, a closer study of the above mentioned principles could help to understand how those abilities could develop prior to the mechanism for more complex social interaction. A next logical step now is to identify mechanisms that can be used to identify agents, from the same agent centred, information theoretic perspective, that share similar goals, and make the agents actively seek out informational benefits by manipulating other agents in the environment. This could bridge the gap from behaviour creation to social interaction, and it would be interesting to study what kind of mechanism would naturally emerge.

References

- [ABD⁺07] Nihat Ay, Nils Bertschinger, Ralf Der, Frank Gttler, and Eckehard Olbrich. Predictive information and explorative behavior of. 2007.
- [CT91] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. John Wiley and Sons, Inc., 1991.
- [DGVW04] Etienne Danchin, Luc-Alain Giraldeau, Thomas J. Valone, and Richard H. Wagner. Public information: From nosy neighbors to cultural evolution. *Science*, 305(5683):487–491, July 2004.
- [KPN07] Alexander S. Klyubin, Daniel Polani, and Chrystopher L. Nehaniv. Representations of space and time in the maximization of information flow in the perception-action loop. *Neural Computation*, 19(9):2387–2432, 2007.
- [Neh99] Chrystopher L. Nehaniv. Meaning for observers and agents. In *IEEE International Symposium on Intelligent Control / Intelligent Systems and Semiotics (ISIC/ISAS'99)*, pages 435–440. IEEE Press, 1999.
- [PMK01] Daniel Polani, Thomas Martinetz, and Jan T. Kim. An information-theoretic approach for the quantification of relevance. In *ECAL '01: Proceedings of the 6th European Conference on Advances in Artificial Life*, pages 704–713, London, UK, 2001. Springer-Verlag.
- [PNMK06] Daniel Polani, Chrystopher L. Nehaniv, Thomas Martinetz, and Jan T. Kim. Relevant information in optimized persistence vs. progeny strategies. In *Artificial Life X : Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems*, pages 337–343. The MIT Press (Bradford Books), August 2006.
- [Sha48] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, July 1948.
- [TPB99] Naftali Tishby, Fernando C. Pereira, and William Bialek. The information bottleneck method. In *Proc. of the 37-th Annual Allerton Conference on Communication, Control and Computing*, pages 368–377, 1999.
- [VVS07] Massimo Vergassola, Emmanuel Villermaux, and Boris I. Shraiman. ‘info-taxis’ as a strategy for searching without gradients. *Nature*, 445(7126):406–409, 2007.