

Local Information Maximisation creates Emergent Flocking Behaviour

Christoph Salge¹ and Daniel Polani¹

¹University of Hertfordshire, Hatfield, UK
c.salge@herts.ac.uk

Abstract

The three boids rules of alignment, separation and cohesion, introduced by Reynolds to recreate flocking behaviour have become a well known standard to create swarm behaviour. In this paper we want to demonstrate how similar flocking behaviour can be created by a local, agent based model, following a principle of information maximisation. The basis for our model is an extension of Vergassola's *infotaxis* model, where agents determine their actions based on the highest expected reduction of entropy. We adapted this approach to a grid world-based search task, and extended the agents abilities so they could not only perform a Bayesian update with information gained from the environment, but also with information gained from other agents. The resulting global flocking behaviour is then analysed in regard to how well it resembles the basic boids rules.

Introduction

Flocking behaviour is a natural phenomenon found in a diverse selection of life forms, such as birds, fish, herd animals and insects. And, as demonstrated by Dyer et al. [8], in specific circumstances even humans exhibit similar behaviour. One of the first models to create this behaviour in a computer simulation is the *boids* steering model, introduced by Reynolds [14]. The model is a prime example of a powerful artificial life idea, namely how local self organisation can create emergent global phenomena. Originally developed to animate the movement of fish and birds for graphical presentation, the boids model has developed into a "de facto" standard for flocking algorithms.

The three basic rules, alignment, separation and cohesion, are agent based and local, so they allow every agent to determine its own actions by itself, using only local data:

- Alignment: steer towards the average heading of local flock mates
- Separation: steer to avoid crowding local flock mates
- Cohesion: steer towards the average position of local flock mates

This model, or variations thereof, are not only the basis for many current flocking and swarm simulations, but are also a powerful example for how simple, local rules can lead to the emergence of complex, life-like properties.

What we want to probe further in this paper is how the global phenomenon of self-organised flocking can be explained; but instead of motivating the individual atomic rules, we intend to challenge the notion that those rules are necessarily atomic. As an alternative, we offer a model where the individual agent's actions, and the resulting global flocking behaviour, is created and motivated by obtaining as much relevant information about the environment as possible. This is an additional result of our previous efforts to extend information theoretic-behaviour generation in general, and in particular the biologically inspired *infotaxis* model by Vergassola et al. [25], to a multiagent system. In the original model the sensor inputs from the environment are used, via a Bayesian Update, to update an internal probabilistic model about a specific location. Actions are chosen based on how much expected information gain they provide for the internal model. In the multiagent model, the actions of other, observable agents are treated with the same Bayesian update, and the resulting agent movement starts to resemble flocking behaviour.

In this paper we are first going to describe our model, and how the single principle of maximal information gain can be used to generate agent behaviour. We shall then demonstrate how information, both from the environment and from other agents, is integrated into the Bayesian model of the agent. The resulting behaviour of those models is then analysed by measuring how well it resembles certain basic characteristics of boids flocking behaviour. We also offer a less formal explanation on how the mechanism of information maximisation leads to flocking behaviour, and how this could be generalised.

Related Work

Information Theory was originally conceived by Shannon [17] to deal with the limits of transatlantic communication; the main focus being the optimal use of a limited commu-

nication channel. But its considerable mathematical versatility, since it can be applied on any system that can be formalised in terms of random variables, also allows for the analysis of a diverse variety of systems in terms of their information theoretic properties and limitations [5].

A recent information theoretic analysis of a boids-like swarm model by [4] demonstrated the ability for information transfer between the flocking agents. Few “informed” agents were capable of steering a swarm. Corresponding results have been observed in the flocking behaviour of human crowds by [8]. If we take a closer look at coordinated systems in nature, it is not surprising that there is a certain degree of mutual information between the organised components. Organisation requires a certain degree of causal dependence, and if we follow the argument of [12], this leads to a certain degree of mutual information between the appropriate variables. Similar conclusion can be drawn for the necessity of information flow, as defined in [2]. The mere presence of some non-vanishing correlation, i.e. nonzero mutual information in nature is, of course, not surprising. However, it is striking that there are many indications that biological organisms tend indeed to operate close to the physical limits for sensory and informational capacities [11, 15]. This can be formulated as an information optimality principle which provides a constructive way to generate behaviours. The use of information theory to model the complexity of cognitive processes [18, 21] has led to systematic approaches to model agent decision making [22, 23, 6] utilizing information theory in a constructive way, beyond the use as a merely analytic tool. To mention a few examples; it has been used to optimize behaviour in a Reinforcement Learning-like context by [20]. Also, for behaviour generation, there is the predictive information maximization [1] which is related to the dynamical systems *homeokinesis* principle by [7].

Another example is the idea of *empowerment* by [9], where an agent tries to act as to maximise the channel capacity between its actuators and sensors which essentially is an optimization of its sensorimotor niche. [3] demonstrates that this principle on its own can already leads to coordinated multiagent behaviour. Note that this shows how, seemingly in opposition to the original philosophy behind information theory which had been designed to carry no semantics, our current work is based on ideas that one is able to distinguish between relevant and non-relevant information.

The *information bottleneck* perspective by [19] demonstrates how the notion of Shannon information can be imbued with relevance, and this can be achieved either through the presence of goals or reward structures [13, 24] or, alternatively, imprinted by the agent-environment interaction itself [10].

This concept of *relevant information*[13] is one we refer to when we later talk about an agent maximising information. *Relevant information* is interpreted here according to the information bottleneck formalism [19, 13]. It quantifies

not *all* information (i.e. possible reduction of uncertainty) in the environment, but only that information which identifies the selection of optimal actions by the agent. Under this perspective, any information in the environment beyond that is ignored.

Information Theory

We consider random variables X which can assume concrete values x . Write $P(X = x)$, or $p(x)$ by abuse of notation, for the probability of X assuming the specific value x . We can now define the entropy $H(X)$ of the random variable X as

$$H(X) = - \sum_x P(X = x) \cdot \log P(X = x) \quad (1)$$

This is often used to describe the uncertainty about the outcome of X . An alternative, equivalent interpretation is to consider $H(X)$ as the average expected “surprise” or the information gained if one was to observe the state of X , if all one knows about X is only its distribution $P(X)$.

The entropy has a number of important properties. Among others, as it is an *a priori* uncertainty, the entropy is larger if the outcomes are more evenly distributed than if the outcomes are more concentrated on a particular value — in other words, concentrated values are easier to predict (and less uncertain) than uniformly spread ones.

Consider now two jointly distributed random variables, X and Y ; then we can calculate the conditional entropy of X given a particular outcome $Y = y$ as:

$$H(X|Y = y) = \sum_x P(X = x|Y = y) \cdot \log P(X = x|Y = y) \quad (2)$$

This can also be generalised to the entropy of X , given the random variable Y in general, and is obtained by averaging over all possible outcomes of Y :

$$H(X|Y) = \sum_y p(y) \cdot H(X|Y = y) \quad (3)$$

This is the entropy of X that remains if Y is known. Finally, consider $H(X)$ and $H(X|Y)$, the entropy of X before and after we learn the state of Y . Thus, their difference is the amount of information we can learn about X by knowing Y . Subtracting one from the other, we get a value called mutual information:

$$I(X; Y) = H(X) - H(X|Y) \quad (4)$$

This is the value we will refer to if we use the term information and it is measured in bits; if one variable is said to have information about another it means that the mutual information between them is non-zero. As the mutual information is symmetrical ([5]), this works both ways, so one variable A contains as much information about B , as B does about A .

Importantly, note that this original notion of information does not include any semantics and only depends on the joint

distribution of X and Y . Therefore, in calculating the entropy and the mutual information measures, the labels of the values of the variables are not relevant.

In the specific model described in the next section, we will maximise the information in respect to a specific location of a resource, but the model is entirely general and the information could correspond to any kind of information about the state of the agent’s environment needed for the agent to increase its performance.

Experimental Model

Scenario

We consider a model consisting of agents situated in a torus-shaped grid world of size $n \times m$ with periodic boundary conditions. Each location in this world is in the set $\mathcal{W} = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. There is one single location of interest F^* , defined also over the set \mathcal{W} . To contextualise, we will call the location the *food source*, but one can interpret it as any other relevant location information, such as position of shelter or mates. The goal of the agents is to determine (not reach) this location in the shortest possible time. The agents’ initial location, and the location of the food are randomly generated at the start of the simulation, and each time step an agent can execute a move action which moves it one cell up, down, left or right. The agent then gets new sensor inputs; it is able to see the state of the world in all cells not more than r cells away from it. Its sensor signal for each cell is a two-state random variable that indicates either that those cells are empty or that they contain the (here unique) food source. After this observation, the agent decides where to move next. This behaviour is repeated until the agent finds the food.

Once the agent finds the food, the agent disappears. An agent that has disappeared does not block other agents, cannot be observed, and its behaviour is not taken into account for the statistical measurements. Note that the food source itself is unaffected from agents finding it.

The above scenario determines the basic properties of our setting. Now, as we are interested in flocking behaviour, for an effective evaluation, the simulation will be run continuously, so the agents have time to form a swarm. Thus, instead of reinitializing the simulation every time one or all agents find the food source, at each time step there is a 3 % chance that the food will be randomly relocated. In this case, all agents’ internal model is reset, so they start a new search. Those agents which have disappeared because they found the food will also be put back into the world in the location they previously disappeared from. The purpose of this is to allow swarms that have already formed to continue their coordinated movement.

Agent Behaviour

In our model, the agents determine their actions by using an internal probability distribution F , which stores informa-

tion about the world. This internal distribution implements a Bayesian model for the location of the food source. More precisely F is also defined over \mathcal{W} , and $P(F = f)$ corresponds to the probability of the food source being in location f , given the agent’s current information.

Initially, all cells have the same probability of $\forall f \in \mathcal{W} : p(f) = 1/(n \cdot m)$, since the agent has no information about the location f . However, as the agent moves around, it can observe different locations in \mathcal{W} , and discovers that some locations are either empty or contain the food source. If f contains the food, then $p(f) = 1$. If f is empty, then $p(f) = 0$.

In both cases the probabilities of the other locations are normalised accordingly, so the sum of probabilities is always one. This operation is functionally identical to actually performing a Bayesian update with the observable environmental random variables, namely, the food state of the cells within the agent’s sensor range.

The remaining uncertainty of the agent about the location of the food source is reflected by the internal probability distribution and can be measured in terms of entropy $H(F)$, where F is the agent-internal random variable corresponding to the expected position of the food.

Infotaxis Search

To generate the agent’s behaviour, we adopt a greedy information gain-maximisation algorithm, called *Infotaxis* by [25]. Infotaxis was shown to provide a biologically plausible principle as to how a moth could use the very sparse information provided by their olfactory sensors to determine the source of pheromones inside a wide area. The main idea is to act in a way that increases the expected gain in information at each time step. We adapted the infotaxis approach for our discrete grid world scenario.

Infotaxis behaviour is generated by the following steps:

1. Determine which action a will likely lead to the largest reduction in entropy $H(F)$, the uncertainty regarding the position of the food source.
2. Take action a and update F with the resulting sensor input.
3. If $H(F) > 0$, then repeat from step 1.

In step 1 the agent has to determine the likely reduction of entropy based on F , the agent’s current “knowledge” about F^* .

Depending on the position $w \in \mathcal{W}$ of the agent, there is a set $\mathcal{S} \subseteq \mathcal{W}$ of the locations that are visible to the sensor of the agent. The visible locations are those within the agent’s sensor range, meaning they are r or less cells away from the agent’s position. If the agent, starting from the current position, takes the action a from its set of available actions \mathcal{A} , it will enter a new state w_a . In this new state the agent can now sense a new set of locations, denoted by \mathcal{S}_a .

To calculate the expected entropy reduction of action a , $\Delta H(a)$, two cases have to be considered. In the first case, the actual location of the food source $f \in \mathcal{W}$ would be inside the newly observed set of positions \mathcal{S}_a , inside the sensor range after the action a was taken by the agent. The agent assumes that this occurs with the probability of

$$P(f \in \mathcal{S}_a) = \sum_{f \in \mathcal{S}_a} P(F = f) \quad (5)$$

in reference to the agent internal model F . In this case the agent's uncertainty after carrying out action a , $H(F_a)$ would be reduced to zero, and the reduction of entropy would be the difference $H(F) - H(F_a) = H(F)$.

In the other case, 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 for F , called F_a . According to Bayes' rule, $P(F_a = f) = 0$ for all $f \in \mathcal{S}_a$, the resulting probability for all observed, empty locations to contain the food source is zero, and the remaining locations are normalized accordingly by:

$$\forall (f \notin \mathcal{S}_a) : P(F_a = f) := \frac{P(F = f)}{\sum_{w \notin \mathcal{S}_a} P(F = w)} \quad (6)$$

This divides the remaining non-zero probabilities, by the sum of their probabilities, normalizing the overall sum of all probabilities to 1. This updated version of F_a can then be used to calculate the reduction of entropy in the second case, which is given by the difference $H(F) - H(F_a)$. If we put all this together, the expected reduction of entropy for taking action a is:

$$\Delta H(a) = P(F \in \mathcal{S}_a) \cdot H(F) + P(F \notin \mathcal{S}_a) \cdot (H(F) - H(F_a)) \quad (7)$$

To summarize, each step the agent selects the action a that maximises $\Delta H(a)$. If several actions lead to the same expected entropy reduction, the agent selects one of them at random. The sensors are then updated as described above, and this behaviour is repeated until the food source is located. Essentially, this behaviour implements a version of Vergassola et al.'s *infotaxis* search and we will refer to it as such in the subsequent text.

Social Bayesian Update

Earlier studies of single agent infotaxis behaviour in [16] demonstrated that the agent's actions contain information about the food source location. If we look at Fig.2, we see how the probability of the food source location is distributed conditioned on an agent moving north. More importantly, every agent which has to take in (a minimum amount of) relevant information to attain a certain performance level also *must necessarily* encode at least that amount of information

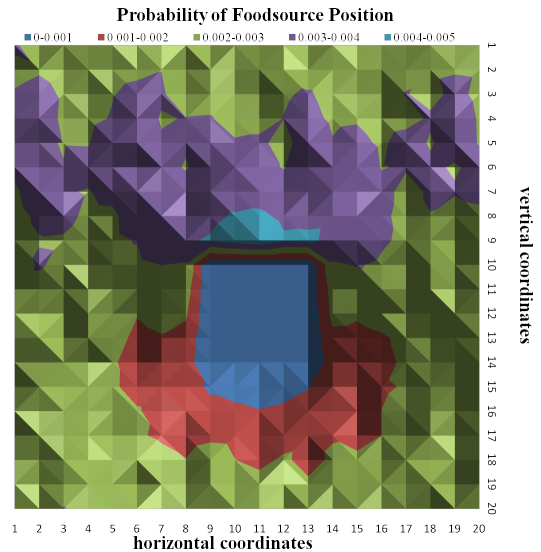


Figure 1: Graph showing $P(F|A = north)$, the probability distribution of F , the food source position, given a specific agent movement (in this case north). The data was obtained from 10000 single agent simulations in a 20×20 grid world, agent position is (11, 11). Note that there is a peak north of the agent, meaning that it is more likely for the food source to be directly north of the agent when it moves north.

in its actions, and this is the case even if it does not have an explicit intention to communicate. This *digested information*, as discussed in [16], has several properties which are interesting for an observing agent with similar goals:

1. Actions *must* contain relevant information, even if the agent does not want to communicate
2. Better agent performance requires more, or the same amount of relevant information
3. The actions of an agent are likely to exhibit a higher density of relevant information than other parts of the environment
4. The actions of an agent might contain information that is not available in the current space or time.

From these properties it follows that a reasonable next step in our information maximisation model would be for the agent to use this digested information and incorporate it in their internal probability distribution. We extend the model so the agent can now, for all cells in its sensor range, detect whether one or more agents are in that cell and where they came from. So, the four new sensor states for each cell are agent that moved in from the north, ... south, ... east, ... west. Each observed move will lead to an adjustment of the assumed internal probability distribution, using a similar

form of Bayesian update already used to integrate the information from the environment. This adjustment of probabilities can be comfortably integrated into our existing infotaxis search.

Note that for the now 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. But they only use this ability if they accidentally encounter another agent. They do not deliberately seek out other agents.

Bayesian Update

Let F denote the agent’s current internal probability model for the location of the food source F^* , and a the state of the random variable A that encodes the last move action of another agent it’s observing. The agent then use Bayes’ Theorem to update the probability distribution of F , with the observed action a .

What the agent is interested in is the probability of the food source to be in a specific location, given the evidence of another agent’s action and relative position $P(F^* = w|A = a)$. According to Bayes’ Theorem this is calculated for every potential location f of the environment as:

$$P(F = f|A = a) = \frac{P(A = a|F^* = f)}{P(A = a)} \cdot P(F = f) \quad (8)$$

Whenever an agent encounters one or several agents it uses this formula to adjust its internal probability $P(F = f)$ for every location of $f \in \mathcal{W}$.

- $P(F = f)$, the *a priori* probability, is the internal model of the agent for mapping the probability distribution of F^* , as gained by their own experience so far;
- $P(A = a)$ is the probability of an agent taking the move action a . Rotational symmetry suggests a probability of 1/4 for each action $a \in \{\textit{north}, \textit{west}, \textit{south}, \textit{east}\}$. Measurements in our single agent simulation confirm this. This is a normalisation factor, so the overall sum of probabilities is still one.
- $P(A = a|F^* = f)$ is the probability of another agent performing action a if the food is in position f . Note that the position f in this case will always be calculated in relation to the position of the observed agent. So, the question we are asking is for example “If the food is known to be 3 cells north of the agent, what is the probability of the agent performing move action a ”. We then record all the cases in the past where an agent has been observed 3 cells south of a food source together with the action it took.

To obtain these statistics for the computer simulation, we observed 10000 single 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 according to the “Infotaxis” part of this paper. So, even though all the agents in the infotaxis simulation have the ability to sense other agents and update their internal world models, they still calculate their Bayesian update under the assumption that all others were non-social agents. We used the data obtained from non-social agents to create the statistics for the probabilities $P(A = a)$ and $P(A = a|F^* = f)$.

After the agent updates F , it resumes the previously described infotaxis behaviour to generate its next move action. Note that agents which have successfully located the food stopped moving and were neither perceivable by other agents, nor blocking them. This was done to increase the challenge since it would have been trivial for another agent to infer from seeing another non-moving agent that the food must be within sensor range of that agent. As a result, the agents could not “cheat” by observing any agents which already knew where the food was.

This model, which includes the Bayesian update not only based on environmental variables, but also on other agents they encounter will be called the *Social Bayesian* model. Apart from the update of the internal model before the next infotaxis action is chosen, it is identical to the infotaxis model.

Measurements

While flocking behaviour might be intuitively visible at this point in our model, defining an objective overall measure which quantifies the emergent flocking behaviour seems difficult. Instead, we aimed to measure the immediate effects of behaving according to the boids rules should have. We defined the following three measurements:

Alignment

To quantify the *alignment* of the different agents, we added up all the agents’ movements and took the length of the resulting vector and normalised it. I.e., every agent $x \in \mathcal{X}$ has an associated vector

$$\vec{v}_x \in \{(1, 0), (0, 1), (-1, 0), (0, -1)\} \quad (9)$$

corresponding to the last direction it moved in. The global alignment is then calculated as the length of the sum of all agent’s vectors, divided by the number of agents:

$$\textit{alignment} = \frac{|\sum_{x \in \mathcal{X}} \vec{v}_x|}{|\mathcal{X}|} \quad (10)$$

This results in a value between 1.0 and 0.0. The maximum value is reached when all agents move in the same direction, and the lowest value of 0.0 is attained when the movement of all agents is distributed evenly between those moving north and south, and those moving west and east, respectively. Note again that agents which have found the food are not taken into consideration for this measurement, since

it would be irrelevant to measure how well aligned they are, once they are not moving anywhere.

This measurement are taken for every simulation step, and an average over all simulation steps is calculated for the whole simulation.

Cohesion

To measure *cohesion*, we simply count, for every agent, how many other agents are within the agent’s sensor range for any given time step. This value is then averaged over all agents, and over all time steps, and the result we call the local agent density, or simply density. This value, different from the global alignment, is only taken locally, and reflects how well agents keep other agents within their own sensor range.

Separation

The hardest value to measure is *separation*, since it basically quantifies an objective of what should not happen. To approximate this, we measure how often one agent tries to enter the cell of another agent, and thus colliding with it. In this case, the agent trying to move will simply fail doing so. The resulting number of overall collisions is then divided by the number of time steps, providing an average amount of collisions per round, or simply collisions. This number is of course also dependent on the number of agents in the simulation, but this correlation if not linear, is therefore not normalised with respect to agent number. Thus, one needs to take care to only compare values where similar amounts of agents have been involved. Again, agents who have found the food are not considered for collisions detection.

Results

All measurements were taken in a open ended simulation where the food had a 3 % chance of being moved every time step. When this happens, all agents’ internal models are reset, and those agents who have already found the food are put back into the simulation. The simulations were run for 100,000 time steps, with 20 agents, in a 20×20 torus-shaped grid world. As a baseline for comparison, we also measured those values for a group of agents that chose their actions at random, only stopping if they chanced upon the food source.

	Alignment	Density	Collisions
Random	0.23	1.03	0.72
Infotaxis	0.29	1.33	1.31
Social B. Update	0.39	1.68	0.49

Table 1: Flocking indication measurements taken for three behaviour models. (Random, Infotaxis, Social Bayesian

If we move from the random behaviour to the single agent infotaxis search, we see both the local agent density and the number of collisions increase. Since agents are not yet reacting to each other in the plain infotaxis model, this seems to

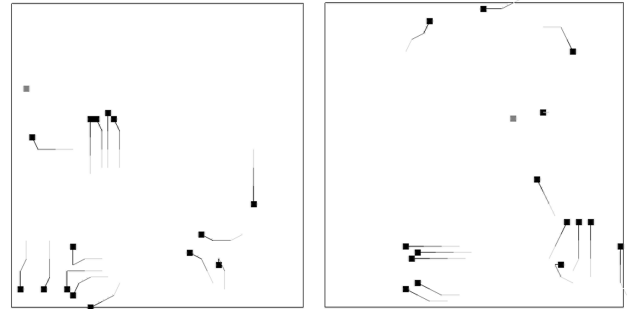


Figure 2: Two screen shots from a social infotaxis simulation with 15 agents, sensor range 5 in a 50 x 50 world. The grey box is the food source, the black boxes are agents. The lines indicate the vector of movement in the last 9 turns, in steps of 3.

be a result of the improved search algorithm. If we measure how long it takes, on average, for a *random* agent to find the food (ca. 450 time steps), and compare it to the time it takes an *infotaxis* agent to find the food (ca. 70 time steps), we see that the infotaxis search has a much better performance, resulting in agents actually finding the food before it changes position. This in turn leads to a local concentration of agents, which is likely to result in increased density and collisions. Note, however, that if we look at the alignment indicator we also see, that even for a group of agents that moves at random the average alignment is not 0.0, but 0.23. This is a statistical effect and not surprising, since it would actually take coordination to ensure that all agent’s movements are always balanced between the different directions.

The interesting comparison is now between the two simpler models and the Social Bayesian Update. In the latter, we see a further increase in alignment, indicating that a high number of agents now move in similar directions during most of the simulation. Keep in mind that to achieve an average of 1.0, all agents would have to move in that same direction, in every turn. We also get a further increase in local agent density, while at the same time the number of collisions is reduced. So while there are now even more agents within the sensor range of each other, the agents manage to collide much less.

Interpretation

We presented a model where the agents’ behaviour is motivated by one single principle or goal, namely to gain as much information about a relevant variable in the environment. To achieve this, the agents take any kind of sensor variable, be it an environmental variable, such as the state of a grid world cell, or the action variables of another agent, and performs a naive Bayesian update on its internal proba-

bilistic model about said relevant variable. The agent's own actions are chosen in regard to which of them provides the greatest expected reduction of entropy, based on the agents' own internal model.

In this section, we would now like to discuss possible explanations on how this information maximisation model leads to the three different rules which create the boids-like flocking behaviour.

Alignment

When an agent is controlled by non-social infotaxis behaviour moves, then its action contains information about the relative position of the food source. If we take a look at an agent moving north (due to rotational symmetry, the actual direction is exchangeable), then the food is more likely to be in a position north of the agent, and less likely to be in a position south of it. This effect, even though the agent does not know where the food is, results from the fact that the agent knows where the food is *not*. As seen in Fig. 2, the probability distribution has its highest peak directly north of the agent, and the minimum of the distribution is in the area south of the agent. Both peaks flatten out the further the cells are away from the agent.

Another agent who observed the first agent move north would perform a Bayesian update on its own assumed probability distribution of the food source. Everything else being equal, this would lead him to "believe" that the food is more likely to be north. The resulting move action would also be to rather move north than in any other direction. A flock of agents, each observing each other, could thereby create a "travelling wave" of high probability immediately outside of their sensor range, driving them all in a similar direction.

The generalised principle here is that an agent 1 observing actions by an agent 2 assumed to have similar goals would lead the original agent 1 to conclude that agent 2 has information that would make such an action reasonable, and in turn, this would make the same action more reasonable for agent 1.

Separation

Whenever agent 1 observes an agent 2 moving in our grid world model, it performs a Bayesian update for the position of the food source. The biggest impact of this update is on the probabilities of the area immediately around agent 2. The cells of the world agent 2 observed in its previous turn are definitely empty, so most of the current area around agent 2 cannot contain any new information for the observer. So while observing another agent is an efficient way to gain information, the immediate environment around that agent becomes informationally unrewarding afterwards. An information-driven search would therefore try to steer away from the immediate area around an observed agent.

In general, if an agent 2 in a specific position reveals information it gets from being in that position to agent 1, then

the more information agent 1 gets from that agent, the less informationally interesting does being in that position become.

Cohesion

In our current model, most of the cohesion seen in our agent groups seems to be a direct result of the high amount of agent alignment. If agents that meet each other move into a similar direction, with similar speed, then they also happen to stay together. In general, it would actually be reasonable to include a further term into the infotaxis mechanism which would account for the amount of information gained from other agents. Following from the "digested information" principle, it is informationally advantageous to keep other agents in sensor range, to be able to use them for a Social Bayesian Update. Seeing another agent, and being able to use the information in its actions increases each agent's expected entropy reduction.

All in all, if we take into account both separation and cohesion, the best solution in terms of information gain seems to be to keep other agents just inside your own maximum sensor range.

Future Work

Since all agents observe each other we suspect there is the distinct possibility that a positive feedback loop can emerge, which detaches itself completely from the environmental information. As an example, an agent might take, for lack of better information, a random action; for example to move up north. Another agent might observe the first, and if it did not know anything apart from the fact that another agent moved north, he also would move north. The first agent in turn might now see the second, observe that the other agent moved north, and take this as good reason to also move north. This vicious feedback circle then continues, reaffirming both agents internal beliefs that "they are doing the reasonable thing". This phenomenon warrants further study, since it could illuminate how in social settings seemingly reasonable assumptions lead to strong "convictions" that are utterly wrong and detached from reality.

Furthermore, it might also be interesting to move the present model from a grid world scenario into a continuous world. This would not only create more realistic animations, but would also be necessary to establish that the observed effects are not just artefacts of the grid world model. The challenge here would be the extension of previously described information theoretic tools to the continuous domain.

Conclusion

We found that information-based social observation mechanisms are able to reproduce several postulated mechanisms of flocking. This is confirmed both by qualitative observation as well as using quantitative measures. Starting with the assumption that every agent needs to obtain some kind

of relevant information from the environment to act intelligently, then most of the arguments follow directly from that. Infotaxis seems to be not only conceptually grounded, but both biologically plausible ([25]), as it leads to behaviour that is very similar to actual moth behaviour, and reasonably efficient for some scenarios; its performance in these scenarios is close to that of an optimal strategy ([16]). Our extension to also include the information offered by other agent's actions is well motivated by the properties of "digested information", and the result is a performance increase beyond the level achievable for a single lone agent ([16]).

At this point in the argument, we already observe emergent flocking behaviour, only motivated by one single utility, the maximum information gain. Note that the relevant information we have been discussing does not necessarily have to be the location of a food source. It could refer to the position of predators, or the location of mates or other types of desirable states, and might lead to similar flocking behaviour via similar mechanisms. The *relevant information* hypothesis can also be applied to a wide variety of agent types, whether birds, fish, herd animals or humans, and could offer a possible *ab initio* explanation for an immediate evolutionary gradient leading to flocking behaviour for a diverse spectrum of organisms.

References

- Ay, N., Bertschinger, N., Der, R., Gttler, F., and Olbrich, E. (2008). Predictive information and explorative behavior of autonomous robots. *European Journal of Physics B*, 63(3):329–339.
- Ay, N. and Polani, D. (2008). Information flows in causal networks. *Advances in Complex Systems*, 11(1):17–41.
- Capdepu, P., Polani, D., and Nehaniv, C. L. (2007). Maximization of potential information flow as a universal utility for collective behaviour. In *Proceedings of the First IEEE Symposium on Artificial Life*, pages 207–213.
- Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. A. (2005). Effective Leadership and Decision-Making in Animal Groups on the Move. *Nature*, 433(7025):513–516.
- Cover, T. M. and Thomas, J. A. (1991). *Elements of Information Theory*. Wiley-Interscience, 99th edition.
- Crutchfield, J. P. and Young, K. (1989). Inferring statistical complexity. *Phys. Rev. Lett.*, 63(2):105–108.
- Der, R., Steinmetz, U., and Pasemann, F. (1999). Homeokinesis — A new principle to back up evolution with learning. In *Proceedings of the International Conference on Computational Intelligence for Modelling Control and Automation (CIMCA'99), Vienna, 17-19 February 1999*.
- Dyer, J., Ioannou, C., Morrell, L., Croft, D., Couzin, I., Waters, D., and Krause, J. (2008). Consensus Decision Making in Human Crowds. *Animal Behaviour*, 75(2):461–470.
- Klyubin, A. S., Polani, D., and Nehaniv, C. L. (2005). Empowerment: a universal agent-centric measure of control. In *Congress on Evolutionary Computation*, pages 128–135.
- Klyubin, A. S., Polani, D., and Nehaniv, C. L. (2007). Representations of space and time in the maximization of information flow in the perception-action loop. *Neural Comput.*, 19:2387–2432.
- Laughlin, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Current Opinion in Neurobiology*, 11(4):475 – 480.
- Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge University Press.
- Polani, D., Nehaniv, C. L., Martinetz, T., and Kim, J. T. (2006). 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).
- Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *SIGGRAPH Comput. Graph.*, 21(4):25–34.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W. (1999). *Spikes*. A Bradford Book. MIT Press.
- Salge, C. and Polani, D. (2011). Digested information as an information theoretic motivation for social interaction. *Journal of Artificial Societies and Social Simulation*, 14(1):5.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423.
- Sporns, O. and Lungarella, M. (2006). Evolving coordinated behavior by maximizing information structure. In *Artificial Life X : Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems*, pages 323–329. International Society for Artificial Life, The MIT Press (Bradford Books).
- Tishby, N., Pereira, F. C., and Bialek, W. (1999). The information bottleneck method. In *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing*, pages 368–377.
- Todorov, E. and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature neuroscience*, 5:11.
- Tononi, G., Sporns, O., and Edelman, G. M. (1994). A Measure for Brain Complexity: Relating Functional Segregation and Integration in the Nervous System. *Proceedings of the National Academy of Science*, 91:5033–5037.
- Touchette, H. and Lloyd, S. (2000). Information-theoretic limits of control. *Phys. Rev. Lett.*, 84(chao-dyn/9905039. 6):1156.
- Touchette, H. and Lloyd, S. (2004). Information-theoretic approach to the study of control systems. *Physica A: Statistical Mechanics and its Applications*, 331(1-2):140 – 172.
- van Dijk, S. G., Polani, D., and Nehaniv, C. L. (2010). What do You Want to do Today? Relevant-Information Bookkeeping in Goal-Oriented Behaviour. In *Artificial Life XII: The 12th International Conference on the Synthesis and Simulation of Living Systems*, pages 176–183.
- Vergassola, M., Villermaux, E., and Shraiman, B. I. (2007). 'infotaxis' as a strategy for searching without gradients. *Nature*, 445(7126):406–409.