Information Gathering Strategies

Eleanor Brush
March 18, 2013

Contents

1 Introduction

There are (at least) two reasons animals might update their opinions according to the opinions of the other animals around them. They might want to be able to learn whatever information the other animals have learned about the environment or it might be advantageous for the animals to come to consensus, for example about where to go.

Previous work on how many neighbors birds have:

- anisotropy in bird flocks suggest they pay attention to seven of their nearest neighbors [**?**]
- statistical mechanical analysis of bird flocks suggest they pay attention to eleven of their nearest neighbors [**?**]
- if birds pay attention to seven of their nearest neighbors they construct an interaction network that optimizes the group's ability to come to a robust consensus decision [**?**]

One explanation is that the birds benefit if they are in a group that can come to a consensus decision. Another explanation is that the birds are optimizing some other measure of performance and in so doing construct a network that enables the group's ability to come to consensus. We are interested in seeing if there are other properties the individual's might be optimizing that leads to robust consensus.

The Question: If individual birds are trying to optimize the speed and reliability with which they can learn an external signal that someone else has picked up, and change the number of birds they pay attention to optimize that performance, how does that affect the interaction network they develop? In particular, does it improve or diminish the group's ability to come to consensus?

2 The Model

There are n nodes in the system each have an opinion x_i and receive information from their neighbors, which they weigh by factors w_{ji} such that $\sum_j w_{ij} = 1$. A node updates its opinions by averaging the difference between its opinion and its neighbors'. In continuous time, the dynamics of this learning process are given by

$$\frac{dx_i}{dt}(t) = \sum_j w_{ij}(x_j(t) - x_i(t)) + \xi_i \tag{1}$$

where ξ_i is random noise. Let L be the Laplacian of the weighted matrix given by the w_{ij} so that

$$L_{ij} = \begin{cases} 1 & , \quad \text{if } i = j \\ -w_{ij} & , \quad \text{else} \end{cases} \tag{2}$$

Then the vector of opinions $\mathbf{x}(t)$ is described by

$$\dot{\mathbf{x}}(t) = -L\mathbf{x}(t) + \xi \tag{3}$$

If the graph described by L is connected, then, in the absence of noise, the system will reach a consensus state in which all nodes have the same opinion, $\mathbf{x}(t) = \alpha\mathbf{1}$.

The learning dynamics change if we introduce a stubborn node r who stops listening to his neighbors, for instance if he learns about something in the environment that is more important than the information coming from his peers. In this case, $x_r(t) = s$ for some constant s and all the other nodes continue to update using their neighbors' opinions as before, so that $\dot{x}_i(t)$ is as in **??** for $i \neq r$. In this case, the consensus opinion will be s, the external signal. This is equivalent to defining a new L^r such that $L^r_{ri} = 0$ for all i. The Laplacian L or L^r affects how quickly consensus can reached and how close to the consensus opinion each node can be when there is noise in the system.

3 Measures of Performance

There are two timescales on which an individual might care about its performance: a transient period during and a steady state period during which noise perturbs the opinions away from the external signal.

To measure the distance from the external signal during a transient period (ignoring noise for the moment), we define, for each node i,

$$\tau_i^{ext} = \sqrt{\int_0^\infty |x_i(t) - s|^2 dt} \tag{4}$$

Let λ^r be second smallest eigenvalue of L^r and v be the associated eigenvector. Then we can approximate $x_i(t) - s \sim e^{-\lambda^r t} v_i$ and

$$\tau_i^{ext} = \sqrt{\int_0^\infty e^{-2\lambda^r t} v_i^2 dt} = \frac{|v_i|}{\sqrt{2\lambda^r}} \tag{5}$$

To measure perturbations away from the external signal due to noise in the steady state, we define, for each node i,

$$\sigma_i^{ext} = \sqrt{\lim_{t\to\infty} E[|x_i(t) - s|^2]} \tag{6}$$

Different nodes in the network might be subject to different sources and magnitudes of noise. We will always assume that $\mathbb{E}[\xi_i] = 0$. Let $d_i = Var[\xi_i]$ and $D = diag(d_1, \ldots, d_n)$. In particular, we will set d_i equal to the number of neighbors influencing node i, so that nodes with more neighbors make noisier estimates. This can be found from the H_2 norm of the matrix L^r (see Appendix and [**?**] for details).

These metrics, τ_i^{ext} and σ_i^{ext}, measure node i's ability to learn the external signal either quickly during the transient period or reliably during the steady state period. To combine them into one performance metric, we introduce a tradeoff parameter, α and measure performance according to $(1-\alpha)\tau_i^{ext} + \alpha\sigma_i^{ext}$.

We also need to quantify the group's ability to come to consensus, which again can be measured on a short or a long timescale. Analogously to our metrics of individual performance, we define

$$\tau^{con} = \frac{1}{\sqrt{2\lambda}} \tag{7}$$

where λ is the second smallest eigenvalue of L. In this case, there is no predefined consensus opinion so we need to find the distance between the vector of opinions and consensus in a way that doesn't depend on what consensus opinion is reached. To do this, we define $y(t)$ to be $x(t) - \langle x(t) \rangle$ and

$$\sigma^{con} = \sqrt{\frac{\lim_{t\to\infty} E[||y(t)||^2]}{n}} \tag{8}$$

We have defined these measures so that τ^{con} is low when the group converges to consensus quickly and σ^{con} is low when the steady state noise around the consensus state is low. Again the metrics for each timescale can be combined into one metric by introducing a tradeoff parameter α and trying to minimize $(1-\alpha)\tau^{con} + \alpha\sigma^{con}$.

4 Learning the external signal versus reaching consensus

Lemma 1 TO BE PROVED: Let L be the Laplacian of the weighted matrix and L^r be L with its r^{th} row set equal to 0. Let the second smallest eigenvalue of L be λ and the smallest eigenvalue of L^r be λ^r. Then $1 \geq \lambda \geq \lambda^r \geq 0$. In particular,

1. $\lambda^r = 0$ if and only if $w_{ir} = 0$ for all i and
2. if $w_{ir} = 1$ for all i then $\lambda^r = 1$.

In particular, this means that $\langle \tau_i^{ext} \rangle_i \geq \tau^{con}$.

Lemma 2 If we define $\delta(t) = x(t) - s\vec{1}$, then $||\delta(t)||^2 \geq ||y(t)||^2$ for all t.

Proof. For the following, let $\langle x \rangle = \frac{\sum_i x_i}{n}$.

$$
\begin{aligned}
y = x - \langle x \rangle &\text{ and } \delta = x - s \\
\Rightarrow ||y||^2 = \sum_i (x_i^2 - 2x_i\langle x \rangle + \langle x \rangle^2) &\text{ and } ||\delta||^2 = \sum_i (x_i^2 - 2x_i s + s^2) \\
\Rightarrow ||y||^2 = \sum_i x_i^2 - 2n\langle x \rangle^2 + n\langle x \rangle^2 &\text{ and } ||\delta||^2 = \sum_i x_i^2 - 2ns\langle x \rangle + ns^2 \\
\Rightarrow ||\delta||^2 - ||y||^2 &= ns^2 - 2ns\langle x \rangle + n\langle x \rangle^2 \\
&= n(\langle x \rangle - s)^2 \\
\Rightarrow ||\delta||^2 - ||y||^2 &\geq 0 \text{ with equality iff } \langle x \rangle = s \\
\Rightarrow ||\delta||^2 &\geq ||y||^2 \text{ with equality iff } \langle x \rangle = s
\end{aligned}
$$

■

Therefore, $\lim_{t\to\infty} E[||\delta(t)||^2] \geq \lim_{t\to\infty} E[||y(t)||^2]$. So that $\sum_{i=1}^n (\sigma_i^{ext})^2 \geq n\sigma^{con}$ and finally this means that $\langle \sigma_i^{ext} \rangle_i \geq \sigma^{con}$. It is therefore not unreasonable to hope that by minimizing τ_i^{ext} and σ_i^{ext} the animals will tend to minimize τ^{con} and σ^{con}.

5 Optimal Strategies

If all n nodes have the same number of neighbors, $2k$, then

$$\tau^{con} = \frac{1}{\sqrt{2 - \frac{2}{k}\sum_{\ell=1}^{k} \cos\left(\frac{2\pi\ell}{m}\right)}} \tag{9}$$

and

$$\sigma^{con} = \sqrt{\frac{2k\sum_{j=1}^{n-1} \frac{1}{2-\frac{2}{k}\sum_{\ell=1}^{k}\cos\left(\frac{2\pi j\ell}{m}\right)}}{n}} \tag{10}$$

For a given tradeoff α, we can find the number of neighbors that optimizes the animals' ability to learn the external signal or the group's ability to come to consensus (Figure **??**). Not surprsingly, as α increases (and long term noise matters more), it becomes better to pay attention to fewer neighbors (but more than 2). We find that the optimal strategy scales sublinearly with network size. Most importantly, the optimal strategy for gathering an external signal is to pay attention to fewer neighbors than the optimal strategy for reaching consensus. This means that optimizing the ability to learn the external signal won't work!

6 Singular Strategies and Evolutionary Dynamics

7 Future Work

1. prove Claim **??** about eigenvalues
2. Information centrality predicts the accuracy of each individual's opinion and eigenvector centrality predicts the speed of each individual's opinion. Is there a measure that predicts performance when both speed and accuracy are valued?
3. prove that spatial periodicity in strategy is optimal for group performance
4. allow nodes to actively optimize their strategies
5. allow for more variation among strategies across nodes and for more complicated network structures
6. multiple signals with different values
7. look at real networks

Appendix

H2 Norm Let y represent the distance between the vector of opinions and consensus, i.e. $y = QxQ^T$, where $Q \in \mathbb{R}^{n-1 \times n}$ is such that $Q\vec{1} = 0$, $QQ^T = I_{n-1}$, and $Q^TQ = I_n - \frac{1}{n}\vec{1} \times \vec{1}^T$. Then

$$\dot{y}(t) = -\overline{L}y(t) + Q\xi(t)$$

where $\overline{L} = QLQ^T$. As shown in [**?**], if $\Sigma_y(t) = \mathbb{E}[y(t)y(t)^T]$, Σ_y at equilibrium solves

$$0 = -\overline{L}\Sigma_y - \Sigma_y\overline{L}^T + D. \tag{11}$$