

# Information Gathering Strategies

Eleanor Brush  
May 1, 2014

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Learning dynamics</b>	<b>1</b>
<b>3</b>	<b>Individual-level measures of performance</b>	<b>2</b>
<b>4</b>	<b>Optimization methods</b>	<b>2</b>
<b>5</b>	<b>ESS</b>	<b>2</b>
5.1	Predation . . . . .	2
5.2	Access to food . . . . .	2
<b>6</b>	<b>Learned optima</b>	<b>2</b>
6.1	Predation . . . . .	2
6.2	Access to food . . . . .	2
<b>7</b>	<b>Comparison to data</b>	<b>3</b>
<b>8</b>	<b>Discussion</b>	<b>3</b>
<b>9</b>	<b>Appendix</b>	<b>3</b>
9.1	Equivalence of dynamics and Hamiltonian . . . . .	3
9.2	Numerically evaluating dynamics . . . . .	3
9.3	Calculating correlations . . . . .	4

## 1 Introduction

Birds in flocks, like animals in many types of social groups, can use their peers to learn about the environment and update their opinions about where to move in that environment. Previous work on starlings has measured the number of other birds an individual pays attention to and shown that this number of “neighbors” leads to a social network that is conducive to the whole flock reaching consensus. It is unclear, however, why an individual would change its behavior, i.e. the number of neighbors it learns from, to achieve this outcome. Individual birds should optimize how well they learn about the environment and should minimize the cognitive and time costs of learning. I find the circumstances in which this individual level optimization can produce the observed group level optimization.

## 2 Learning dynamics

In this model, an individual’s opinion changes with a probability that depends on its social interactions and the strength of the environmental stimulus. Specifically, an individual’s opinion  $\sigma_i$  can be either 1 or  $-1$  and  $\sigma_i$  switches with probability  $w_i(\sigma_i) = \frac{1}{2}(-\sigma_i \sum_j A_{ij}\sigma_j)(1 - \beta_i\sigma_i)$  where  $A_{ij}$  indicates the strength of the connection from  $j$  to  $i$  and  $\beta_i$  indicates the strength of the environmental stimulus perceived by  $i$ . The stochastic dynamics of the opinions can be written down with a master equation [4]. The vector  $\vec{v}(t)$  of expected values,  $q_i(t) = \langle \sigma_i(t) \rangle$ , satisfies the following differential equations: [4]

$$\frac{d\vec{v}}{dt} = M\vec{v} + \vec{\beta}^T(\vec{1} - \vec{v}).$$

We simplify this by writing  $\frac{d\vec{v}}{dt} = A\vec{v} + \vec{\beta}^T \vec{1}$ , where  $A = M - B$  and  $B$  is a diagonal matrix with the values of  $\vec{\beta}$  along the diagonal.

These stochastic dynamics will lead to an equilibrium probability distribution over sets of opinions that agrees with the Boltzmann equilibrium distribution for a system with energy or Hamiltonian given by

$$E(\vec{v}) = -\vec{v}^T A \vec{v} - \vec{\beta}^T \vec{v}$$

### 3 Individual-level measures of performance

The group is made up of  $N$  individuals. Each individual has a strategy that determines how many other individuals he pays attention to,  $n_i = \sum_{j \neq i} \mathbf{I}(A_{ij} \neq 0)$ . For a given number of iterations, we distribute the individuals randomly in space. Each individual then pays attention to its  $n_i$  nearest neighbors. This is meant to capture One individual is chosen to be a receiver; it and all those individuals within a radius  $r$  of the receiver perceive an external signal of strength  $\beta$ . We find the expected opinion of each individual  $q_i$  after a period of time  $T$ . We refer to each of these random iterations as a signaling event.

In the absence of an external signal, neither opinion is correct. When an external signal is presented, some individuals will be able to perceive it ( $\beta_i > 0$ ) and some individuals will not ( $\beta_i = 0$ ). In this case, having an opinion that matches the external signal ( $\sigma_i = 1$ ) is beneficial. How beneficial it is depends on the kind of signal. We consider two types of signals: predators and food. If a predator is present, we assume that whichever individual is least aware of the signal will be predated. We therefore find, over many signaling events, the probability that an individual will be eaten and an individual's fitness is given by the probability of surviving. If the signal indicates the presence of food, on the other hand, fitness may be proportional to how well an individual learns the external signal.

### 4 Optimization methods

To understand what strategies we might expect to find, we are interested in identifying the optimal strategies. We do this in two ways. First, using the framework of adaptive dynamics in finite populations, we identify the evolutionarily stable strategy. This framework assumes that there is a homogeneous population into which a mutant individual tries to invade. If strategies change on a learning rather than an evolutionary timescale, or if there is just a lot of variation in the population, we might expect individuals to try to optimize their strategies in the context of a heterogeneous population. This is our second method of optimization.

## 5 ESS

### 5.1 Predation

### 5.2 Access to food

## 6 Learned optima

### 6.1 Predation

### 6.2 Access to food

## 7 Comparison to data

## 8 Discussion

## 9 Appendix

### 9.1 Equivalence of dynamics and Hamiltonian

**Claim 1** If  $H(\vec{v}) = -\vec{v}^T A \vec{v} - \vec{\beta}^T \vec{v}$  and  $\frac{d\vec{v}}{dt} = A\vec{v} + \vec{\beta}$  then  $\nabla H \cdot \frac{d\vec{v}}{dt} \leq 0$ .

*Proof.*

$$\begin{aligned}
H(\vec{v}) &= -\sum_{ij} A_{ij} v_i v_j - \sum_i \beta_i v_i \\
\Rightarrow \frac{\partial H}{\partial v_i} &= -\sum_{j \neq i} A_{ij} v_j - \sum_{j \neq i} A_{ji} v_j - 2A_{ii} v_i - \beta_i \\
&= -\sum_j (A_{ij} + A_{ji}^T) v_j - \beta_i \\
\Rightarrow \nabla H &= -(A + A^T) \vec{v} - \vec{\beta} \\
\Rightarrow \nabla H \cdot \frac{d\vec{v}}{dt} &= -\vec{v}^T (A + A^T) A \vec{v} - \vec{v}^T (A + A^T) \vec{\beta} - \vec{\beta}^T A \vec{v} - \vec{\beta}^T \vec{\beta} \\
&= -\vec{v}^T A^2 \vec{v} - \vec{v}^T A^T A \vec{v} - \vec{v}^T A \vec{\beta} - 2\vec{v}^T A^T \vec{\beta} - \vec{\beta}^T \vec{\beta}
\end{aligned}$$

If  $A$  is negative semi-definite, then  $A^2$  is positive semi-definite and  $A^T A$  is positive semi-definite. (Consider that  $\vec{v}^T A^T A \vec{v} = (A\vec{v})^T (A\vec{v})$  so that  $\vec{v}^T A^T A \vec{v} = 0$  if  $A\vec{v} = 0$  and  $\vec{v}^T A^T A \vec{v} < 0$  otherwise.) Therefore,  $-\vec{v}^T A^2 \vec{v} \leq 0$  and  $-\vec{v}^T A^T A \vec{v} \leq 0$ . At equilibrium,  $\vec{v}$  will be “close” to  $\vec{\beta}$  so that

$$-\vec{v}^T A \vec{\beta} - 2\vec{v}^T A^T \vec{\beta} \sim -\vec{\beta}^T A \vec{\beta}$$

■

### 9.2 Numerically evaluating dynamics

**Claim 2** Let  $\Phi(t)$  be the fundamental matrix solution to the homogeneous differential equations  $\frac{d\vec{x}}{dt} = A\vec{x}$  so that  $\Phi(t) = V\Lambda(t)$  where the columns of  $V$  are the eigenvectors of  $A$  and

$$\Lambda(t) = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})$$

where  $\{\lambda_i\}$  are the eigenvalues of  $A$ . (Note that if  $\lambda_i = 0$  then  $e^{\lambda_i t} = 1$  for all  $t$ .) Let  $V^\dagger$  be the pseudoinverse of  $V$ . Let  $\vec{c} = V^\dagger \vec{X}_0$ . Let

$$G(t) = \text{diag}(1/\lambda_1(1 - e^{-\lambda_1 t}), \dots, 1/\lambda_n(1 - e^{-\lambda_n t})) V^\dagger \vec{\beta}.$$

(If  $\lambda_i = 0$  then the corresponding entry of the first part of  $G(t)$  will be given by  $t$  rather than the form above.) THEN if we let  $x(t) = \Phi(t)G(t) + \Phi(t)\vec{c}$ ,  $X(t)$  solves the inhomogeneous equations  $\frac{dx}{dt} = Ax + \vec{\beta}$  with initial conditions  $x(0) = \vec{X}_0$ .

*Proof.* It is clear that

$$\frac{d\Phi\vec{c}}{dt} = A\Phi(t)\vec{c} \text{ and } \Phi(0)\vec{c} = \vec{X}_0.$$

Now,

$$\begin{aligned}
\Phi(t)G(t) &= V \begin{pmatrix} e^{\lambda_1 t} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} \frac{1}{\lambda_1}(1 - e^{\lambda_1 t}) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_n}(1 - e^{\lambda_n t}) \end{pmatrix} V^\dagger \vec{\beta} \\
&= V \begin{pmatrix} \frac{1}{\lambda_1}(e^{\lambda_1 t} - 1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_n}(e^{\lambda_n t} - 1) \end{pmatrix} V^\dagger \vec{\beta} \\
\Rightarrow \frac{d\Phi(t)G(t)}{dt} &= V \begin{pmatrix} e^{\lambda_1 t} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n t} \end{pmatrix} \\
&= V \begin{pmatrix} e^{\lambda_1 t} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n t} \end{pmatrix} V^\dagger \vec{\beta} - V V^\dagger \vec{\beta} + \vec{\beta} \\
&= V \begin{pmatrix} e^{\lambda_1 t} - 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n t} - 1 \end{pmatrix} V^\dagger \vec{\beta} + \vec{\beta} \\
&= V \begin{pmatrix} \lambda_1 e^{\lambda_1 t} \cdot \frac{1}{\lambda_1}(1 - e^{-\lambda_1 t}) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n e^{\lambda_n t} \cdot \frac{1}{\lambda_n}(1 - e^{-\lambda_n t}) \end{pmatrix} V^\dagger \vec{\beta} + \vec{\beta} \\
&= AV\Lambda(t) \begin{pmatrix} \frac{1}{\lambda_1}(1 - e^{-\lambda_1 t}) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_n}(1 - e^{-\lambda_n t}) \end{pmatrix} V^\dagger \vec{\beta} + \vec{\beta} \\
&= AV\Lambda(t)G(t) + \vec{\beta} \\
&= A\Phi(t)G(t) + \vec{\beta}
\end{aligned}$$

Therefore  $\frac{dx}{dt} = A\Phi(t)G(t) + A\Phi(t)\vec{c} + \vec{\beta} = Ax(t) + \vec{\beta}$  and since  $G(0) = \vec{0}$ ,  $x(0) = \vec{X}_0$ . ■

**9.3 Calculating correlations** The following is derived from the work in [2].

$$H(\vec{v}) = -\vec{v}^T A \vec{v} - \vec{\beta}^T \vec{v}$$

Define  $V = \sum_k v_k / N$  and  $\vec{y} = (\vec{v} - V\vec{1})/V$ . Then  $\vec{v} = V(\vec{1} + \vec{y})$ .

$$\begin{aligned}
H(\vec{y}, V) &= -V^2(\vec{1}^T + \vec{y}^T)A(\vec{1} + \vec{y}) - \vec{\beta}^T V(\vec{1} + \vec{y}) \\
&= -V^2(\vec{1}^T + \vec{y}^T)A\vec{y} - V\vec{\beta}^T \vec{y} - V\vec{\beta}^T \vec{1} \text{ since } A\vec{1} = \vec{0} \\
&= -V^2\vec{y}^T A\vec{y} + (-V^2\vec{1}^T A - V\vec{\beta}^T)\vec{y} - V\vec{\beta}^T \vec{1}
\end{aligned}$$

Define  $Q \in M_{N-1, N}$  that rotates an  $N$ -vector away from the consensus vector and define  $\vec{z} = Q\vec{y}$  so that  $\vec{y} = Q^T \vec{z}$  since  $\sum_k y_k = 0$ .

$$\begin{aligned}
H(\vec{z}, V) &= -V^2\vec{z}^T Q A Q^T \vec{z} + (-V^2\vec{1}^T A Q^T - V\vec{\beta}^T Q^T)\vec{z} - V\vec{\beta}^T \vec{1} \\
&= V^2\vec{z}^T P \vec{z} + (-V^2\vec{\sigma}_1^T - V\vec{\sigma}_2^T)\vec{z} - V\vec{\beta}^T \vec{1} \\
\Rightarrow P(\vec{z}, V) &= \frac{1}{\mathcal{Z}} \exp \left( -V^2\vec{z}^T P \vec{z} + (V^2\vec{\sigma}_1^T + V\vec{\sigma}_2^T)\vec{z} + V\vec{\beta}^T \vec{1} \right)
\end{aligned}$$

where  $P = -QAQ^T$ ,  $\vec{\sigma}_1 = QA^T \vec{1}$ , and  $\vec{\sigma}_2 = Q\vec{\beta}$ .

**Fact 3**

$$\begin{aligned}
\int \exp \left( -\frac{1}{2} \bar{z}^T P \bar{z} + \bar{\sigma}^T \bar{z} \right) d^{N-1} z &= \sqrt{\frac{(2\pi)^{N-1}}{\det P}} \exp \left( -\frac{1}{2} \bar{\sigma}^T P^{-1} \bar{\sigma} \right) \\
\Rightarrow \int \exp \left( -\bar{z}^T P \bar{z} + \bar{\sigma}^T \bar{z} \right) d^{N-1} z &= \sqrt{\frac{(2\pi)^{N-1}}{2^{N-1} \det P}} \exp \left( -\frac{1}{4} \bar{\sigma}^T P^{-1} \bar{\sigma} \right) \\
&= \sqrt{\pi^{N-1}} \sqrt{\frac{1}{\det P}} \exp \left( -\frac{1}{4} \bar{\sigma}^T P^{-1} \bar{\sigma} \right)
\end{aligned}$$

$$\begin{aligned}
\mathcal{Z} &= \int_{-\infty}^{\infty} \int_{\mathbb{R}^{N-1}} \exp \left( V^2 \bar{z}^T P \bar{z} + (V^2 \bar{\sigma}_1^T + V \bar{\sigma}_2^T) \bar{z} + V \bar{\beta}^T \bar{\Gamma} \right) d^N z dV \\
&= \int_{-\infty}^{\infty} \exp(V \bar{\beta}^T \bar{\Gamma}) \int_{\mathbb{R}^{N-1}} \exp \left( V^2 \bar{z}^T P \bar{z} + (V^2 \bar{\sigma}_1^T + V \bar{\sigma}_2^T) \bar{z} \right) d^N z dV \\
&= \sqrt{\pi^{N-1}} \int_{-\infty}^{\infty} \exp(V \bar{\beta}^T \bar{\Gamma}) \times \left( \sqrt{\frac{1}{V^{2(N-1)} \det P}} \right) \exp \left( -\frac{1}{4} \bar{\sigma}_1^T P^{-1} \bar{\sigma}_1 \right) \exp \left( -\frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) dV \\
&= \sqrt{\pi^{N-1}} \sqrt{\frac{1}{\det P P}} \exp \left( -\frac{1}{4} \bar{\sigma}_1^T P^{-1} \bar{\sigma}_1 \right) \int_{-\infty}^{\infty} \exp(V \bar{\beta}^T \bar{\Gamma}) \times \left( \sqrt{\frac{1}{V^{2(N-1)}}} \right) \exp \left( -\frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) dV \\
&= \sqrt{\pi^{N-1}} \sqrt{\frac{1}{\det P}} \exp \left( -\frac{1}{4} \bar{\sigma}_1^T P^{-1} \bar{\sigma}_1 \right) \int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp \left( V \bar{\beta}^T \bar{\Gamma} - \frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) dV \\
&= \sqrt{\pi^{N-1}} \sqrt{\frac{1}{\Pi_a \lambda_a}} \exp \left( -\frac{1}{4} \bar{\sigma}_1^T P^{-1} \bar{\sigma}_1 \right) \int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp \left( V \bar{\beta}^T \bar{\Gamma} - \frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) dV
\end{aligned}$$

where  $\lambda_a$  are the eigenvalues of  $P$

$$\Rightarrow -\log(\mathcal{Z}) = \text{constants} + \frac{1}{2} \sum_a \log(\lambda_a) + \frac{1}{4} \bar{\sigma}_1^T P^{-1} \bar{\sigma}_1 - \log(f(P))$$

$$\text{where } f(P) = \int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp \left( V \bar{\beta}^T \bar{\Gamma} - \frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) dV.$$

$$\begin{aligned}
\Rightarrow \frac{\partial -\log(\mathcal{Z})}{\partial P_{ij}} &= \frac{1}{2} \sum_a \frac{w_i w_j}{\lambda_a} + \frac{1}{4} \sum_{kl} \sigma_{1k} P_{ki}^{-1} P_{jl}^{-1} \sigma_{1l} - \frac{\partial \log f}{\partial P_{ij}} \\
\Rightarrow \frac{\partial -\log(\mathcal{Z})}{\partial P_{ij}} &= \frac{1}{2} \sum_a \frac{w_i w_j}{\lambda_a} + \frac{1}{4} \sum_{kl} \sigma_{1k} P^{-1} k i P_{jl}^{-1} \sigma_{1l} - \frac{1}{f} \frac{\partial f}{\partial P_{ij}}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial f}{\partial P_{ij}} &= \int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp \left( V \bar{\beta}^T \bar{\Gamma} - \frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) \left( -\frac{1}{4V} \sum_{kl} \sigma_{2k} P_{ki}^{-1} P_{jl}^{-1} \sigma_{2l} \right) dV \\
&= \left( -\frac{1}{4} \sum_{kl} \sigma_{2k} P_{ki}^{-1} P_{jl}^{-1} \sigma_{2l} \right) \int_{-\infty}^{\infty} \frac{1}{V^{(N)}} \exp \left( V \bar{\beta}^T \bar{\Gamma} - \frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) dV
\end{aligned}$$

$$\Rightarrow \langle z_i z_j \rangle = \frac{1}{2} \sum_a \frac{w_i w_j}{\lambda_a} + \frac{1}{4} \sum_{kl} \sigma_{1k} P^{-1} k i P_{jl}^{-1} \sigma_{1l} + \left( \frac{1}{4} \sum_{kl} \sigma_{2k} P_{ki}^{-1} P_{jl}^{-1} \sigma_{2l} \right) \frac{\int_{-\infty}^{\infty} \frac{1}{V^{(N)}} \exp \left( V \bar{\beta}^T \bar{\Gamma} - \frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) dV}{\int_{-\infty}^{\infty} \frac{1}{V^{(N-1)}} \exp \left( V \bar{\beta}^T \bar{\Gamma} - \frac{1}{4V} \bar{\sigma}_2^T P^{-1} \bar{\sigma}_2 \right) dV}$$

## References

- [1] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior

depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Science*, 105(4):1232–1237, January 2008.

- [2] William Bialek, Andrea Cavagna, Irene Giardina, Thierry Mora, Oliver Pohl, Edmondo Silvestri, Massimiliano Viale, and Aleksandra Walczak. Social interactions dominate speed control in driving natural flocks toward criticality. *arXiv*, 2013.
- [3] William Bialek, Andrea Cavagna, Irene Giardina, Thierry Mora, Edmondo Silvestri, Massimiliano Viale, and Aleksandra M. Walczak. Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Science*, 109(13):4786–4791, March 2012.
- [4] Roy J. Glauber. Time dependent statistics of the ising model. *Journal of Mathematical Physics*, 4(2):294–307, 1963.
- [5] Petter Holme and Gourab Ghoshal. Dynamics of networking agents competing for high centrality and low d. *Phys. Rev. Lett.*, 96(4), March 2006.
- [6] Carl Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 2001.
- [7] Martin Nowak. *Evolutionary Dynamics*. Belknap / Harvard, 2006.
- [8] George F. Young, Luca Scardovi, Andrea Cavagna, Irene Giardina, and Naomi E. Leonard. Starling flock networks manage uncertainty in consensus at low cost. *PLOS Computational Biology*, 9, 2013.
- [9] George Forrest Young, Luca Scardovi, and Naomi Ehrlich Leonard. Robustness of noisy consensus dynamics with directed communication. *Proceedings of the American Control Conference*, 2010.