Biodiversity, Entropy and Thermodynamics

John Baez
http://math.ucr.edu/home/baez/bio_info/

October 29, 2014
Biological and Bio-Inspired Information Theory
BIRS

Please apply here to attend the workshop I'm running with John Harte and Marc Harper:

- **Information and Entropy in Biological Systems**, National Institute for Mathematical and Biological Synthesis, Knoxville, Tennesee, Wednesday-Friday, 8-10 April 2015.

  http://www.nimbios.org/workshops/WS_entropy

Shannon entropy

$$S(p) = -\sum_{i=1}^{n} p_i \ln(p_i)$$

is fundamental to thermodynamics and information theory. But it's also used to measure biodiversity, where p_i is the probability that a randomly chosen organism is of the ith species.

Is this a coincidence? *No!*

- In thermodynamics, the entropy of system is the expected amount of information we gain by learning its precise state.
- In biodiversity studies, the entropy of an ecosystem is the expected amount of information we gain about an organism by learning its species.

If biodiversity is a form of entropy, shouldn't it always increase?

No:

- The entropy contained in biodiversity is just a tiny portion of the total entropy of an ecosystem. Entropy can increase while biodiversity decreases.
- Entropy can decrease on Earth while total entropy increases: incoming sunlight gets re-emitted as infrared.

However, something *like* the Second Law of Thermodynamics holds whenever an ecosystem has a 'dominant' mixture of species, which can invade all others.

Let

$$P = (P_1, \ldots, P_n)$$

be the vector of populations of n different self-replicating entities, which we'll call **species** of **organisms**.

The probability that a randomly chosen organism belongs to the ith species is

$$p_i = \frac{P_i}{\sum_j P_j}$$

We can think of this probability distibution as a 'hypothesis' and its change with time as a 'learning process'. Natural selection is analogous to Bayesian updating.

Suppose the **replicator equation** holds:

$$\frac{dP_i}{dt} = F_i(P_1, \ldots, P_n)\; P_i$$

Here the population P_i changes at a rate proportional to P_i, but the 'constant of proportionality' can be any smooth function of the populations of all the species.

We call $F_i = F_i(P_1, \ldots, P_n)$ the **fitness** of the ith species.

Let's see how the probabilities change:

$$
\begin{aligned}
\frac{dp_i}{dt} &= \frac{d}{dt}\, \frac{P_i}{\sum_j P_j} = \frac{\dot{P}_i \sum_j P_j \quad - \quad P_i\ \sum_j \dot{P}_j}{\left(\sum_j P_j\right)^2} \\
&= \frac{F_i P_i}{\sum_j P_j} \quad - \quad \frac{P_i\ \sum_j F_j P_j}{\left(\sum_j P_j\right)^2} \\
&= F_i p_i - \left(\sum_j F_j p_j\right) p_i
\end{aligned}
$$

$\overline{F} = \sum_j F_j p_j$ is the **mean fitness** of the population. So:

$$
\frac{dp_i}{dt} = \left(F_i - \overline{F}\right) p_i
$$

$$\frac{dp_i}{dt} = \left(F_i - \overline{F}\right) p_i$$

For your market share to increase, you don't need to be good: *you just need to be better than average.*

If $f_i = F_i - \overline{F}$ is the **excess fitness** of the ith species, we obtain the **Lotka–Volterra equation**:

$$\frac{dp_i}{dt} = f_i p_i$$

Remember, f_i is a function of the population $P = (P_1, \ldots, P_n)$.

Does the Lotka–Volterra equation implies a version of the Second Law of Thermodynamics?

Under some conditions, yes! But it involves *relative* entropy.

Let p and q be a two probability distributions. The **information of *q* relative to *p***, or **Kullback–Leibler divergence**, is

$$I(q,p) = \sum_i q_i \ln\left(\frac{q_i}{p_i}\right)$$

This is the amount of information *left to learn* if p is our current hypothesis and the 'true' probability distribution describing a situation is q.

In Bayesian language, p is our 'prior'.

Let's see how the relative information changes as $p(t)$ evolves, holding q fixed:

$$\begin{aligned}
\frac{d}{dt} I(q, p(t)) &= \frac{d}{dt} \sum_i q_i \ln\left(\frac{q_i}{p_i(t)}\right) \\
&= -\frac{d}{dt} \sum_i q_i \ \ln(p_i(t)) \\
&= -\sum_i q_i \frac{\dot{p}_i}{p_i} \\
&= -\sum_i q_i f_i
\end{aligned}$$

using the Lotka–Volterra equation in the last step.

When is

$$\frac{d}{dt} I(q, p(t)) \;=\; -\sum_i q_i f_i$$

actually ≤ 0, so the 'information left to learn' decreases?

$$\sum_i q_i f_i \;=\; \sum_i q_i\, f_i(P_1, \ldots, P_n)$$

is the average excess fitness for a small 'invader' population with distribution q_i, given that the overall population is $P = (P_1, \ldots, P_n)$. It says how much fitter, on average, the invaders are. If this is ≥ 0 for all choices of P, we call q a **dominant** distribution.

And in this case we get the Second Law: $\dfrac{d}{dt} I(q, p(t)) \leq 0$.

So: if there is a dominant distribution — a probability distribution of species whose mean fitness is at least as great as that of any population it finds itself amidst — then we get a version of the Second Law. *As time passes, the information* **any** *population has 'left to learn' always decreases.*

Reality is more complicated, but this is a nice start.

This result was shown by Akin and Losert in 1984. For more, see:

- Marc Harper, The replicator equation as an inference dynamic, arXiv:0911.1763.
- Marc Harper, Information geometry and evolutionary game theory, arXiv:0911.1383.

and newer papers by Harper.