Measuring Autonomy and Emergence via Granger Causality

Anil K. Seth*
University of Sussex

Keywords
Autonomy, emergence, Granger causality, consciousness, flocking, evolution

Abstract Concepts of emergence and autonomy are central to artificial life and related cognitive and behavioral sciences. However, quantitative and easy-to-apply measures of these phenomena are mostly lacking. Here, I describe quantitative and practicable measures for both autonomy and emergence, based on the framework of multivariate autoregression and specifically Granger causality. *G-autonomy* measures the extent to which knowing the past of a variable helps predict its future, as compared to predictions based on past states of external (environmental) variables. *G-emergence* measures the extent to which a process is both *dependent upon* and *autonomous from* its underlying causal factors. These measures are validated by application to agent-based models of predation (for autonomy) and flocking (for emergence). In the former, evolutionary adaptation enhances autonomy; the latter model illustrates not only emergence but also downward causation. I end with a discussion of relations among autonomy, emergence, and consciousness.

1 Introduction

Concepts of emergence and autonomy are central to artificial life and related cognitive and behavioral sciences. However, quantitative and easy-to-apply measures of these phenomena are mostly lacking. This is unfortunate, because the ability to measure a phenomenon is an essential step toward its effective scientific description [13]. In this article, I introduce quantitative and practicable measures for both autonomy and emergence, based on the framework of multivariate autoregression and specifically Granger causality [19]. These measures are validated by application to simple agent-based models of predator-prey behavior (for autonomy) and flocking (for emergence).

1.1 Autonomy

Autonomy has a wide variety of meanings. A non-exhaustive list might include the ability to learn, the ability to self-recharge, and the ability to operate without power cables or without teleoperation. A more abstract definition of autonomy as "organizational closure" is prominent in the work of Varela (see, e.g., [47]). Here, motivated by practical applicability, I adopt a simple conception of autonomy as the degree of self-determination of a system [10]. Following Bertschinger and colleagues [9], I amplify this concept in two ways: (i) An autonomous system should not be fully determined by its environment, and (ii) a random system should not have a high autonomy value. I introduce a quantitative measure, *G-autonomy*, which operationalizes this notion. Broadly, a G-autonomous variable is one for which prediction of its

* School of Informatics and Sackler Centre for Consciousness Science, University of Sussex, Brighton, BN1 9QJ, UK. E-mail: a.k.seth@sussex.ac.uk, www.anilseth.com

future evolution is enhanced by considering its own past states, as compared to predictions based on past states of a set of external variables. This measure is based on *Granger causality*, a statistical interpretation of causality according to which a signal A causes a signal B if information in the past of A helps predict the future of B, over and above predictions based on the past of B alone [19, 37].

Having a practically applicable measure of autonomy is useful for gaining insight into mechanisms underlying apparently autonomous behavior in organisms, as well as into selective pressures that can lead to increases or decreases in autonomy. For example, in the simulation model described in Section 3, evolutionary adaptation increases autonomy in a predator-prey system. Autonomy measures are not limited to analysis of agent behavior; one can measure the autonomy of any variable that varies over time with respect to other variables. For example, it is possible to measure the autonomy of the activity of brain regions with respect to activity in other regions, and to assess the task and state dependence of these relations.

1.2 Emergence

The concept of emergence has a long history of philosophical discussion [11, 27]. Its essence, however, is simple enough: An emergent property is in a qualitative sense "more than the sum" of its component parts. Emergent properties appear to arise in complex systems of all kinds: biological, cognitive, social, and technological. Broadly speaking, artificial life and complexity science focus on explaining phenomena that seem to involve emergence, and models constructed under these auspices are often described as emergent [6–8].

Following Bedau, emergence can be differentiated into three categories: strong, weak, and nominal [6, 7] (see also [5, 46]). The least controversial is nominal emergence, which is simply the notion of a kind of property that can be possessed by macro-level objects or processes but not by their micro-level constituents. For example, a circle is nominally emergent from the set of points from which it is constructed.

Most controversial is the notion of strong emergence, which involves two related claims. First, a macro-level property is *in principle* not deducible from micro-level observations. Second, macro-level properties have irreducible causal powers. The first claim rejects mechanistic explanations altogether, apparently calling a halt to scientific advance in the absence of new principles of nature [12]. The second involves the problematic notion of *downward causation*. Downward causation is problematic firstly because it contravenes the plausible doctrine that "the macro is the way it is in virtue of how things are at the micro," an idea that has been expressed variously as *causal fundamentalism* [25] or *supervenience* [27]. A second challenge raised by downward causation is that of resolving conflicts between micro-level and macro-level causes [7].

In between strong emergence and nominal emergence lies the useful notion of weak emergence [6–8], according to which a macro-level property is derived from the interaction of micro-level components but in complicated ways such that the macro-level property has no simple micro-level explanation. In contrast to strong emergence, weakly emergent properties are in principle deducible from micro-level components, and in contrast to nominal emergence, the micro-to-macro inferential pathways must be nontrivial. According to Bedau, weakly emergent macro-level properties are *ontologically dependent* on and reducible to micro-level causal factors, but at the same time they are *epistemologically irreducible* due to the complexity of the micro-to-macro inferential pathways. More specifically, on Bedau's view a weakly emergent property is one that has an *incompressible explanation*. An incompressible explanation is one which requires "crawling the causal web" of a system, that is, aggregating and iterating all micro-level interactions over time.[1] According to Bedau, weak emergence can occur in degrees; however, these degrees are qualitative rather than quantitative (e.g., properties with actually incompressible explanations, properties with compressible explanations that cannot be understood, and so on). In this view, there remains no way to *measure* quantitatively the degree of weak emergence exhibited by a system.

1 In previous work Bedau has expressed the same idea with different terms, namely that a weakly emergent property is "underivable except by simulation" [7]. His new terminology shifts emphasis from derivation to explanation but is nonetheless equivalent. The present notion of G-emergence is based neither on simulation nor on explanation, but on prediction.

Here, I propose a continuous version of weak emergence that leverages both Granger causality and the measure of autonomy, G-autonomy, introduced above. I describe a new quantitative measure, *G-emergence*, which operationalizes the notion that a weakly emergent process is simultaneously autonomous from and dependent upon its underlying causal factors. I show that this measure behaves appropriately in a simulation model reflecting a canonical instance of emergence—flocking behavior—and I show how G-emergence can support a practically useful and ontologically innocent means of ascribing downward causality from macro levels to micro levels.

1.3 Organization

The remainder of this article is arranged as follows. Section 2 provides the mathematical detail underpinning Granger causality, G-autonomy, and G-emergence. Sections 3 and 4 describe simulation models of predator-prey interaction and flocking behavior, which serve to validate G-autonomy and G-emergence, respectively. Section 5 discusses wider issues raised by having quantitative measures of autonomy and emergence, including concepts of downward causality and the relation between emergence and consciousness.

2 Theoretical Framework

2.1 Granger Causality

In 1969 Granger introduced the idea of Granger causality (G-causality) as a formalization, in terms of linear regression modeling, of Wiener's intuition that Y causes X if knowing Y helps predict the future of X [19, 37]. According to G-causality, Y causes X if the inclusion of past observations of Y reduces the prediction error of X in a linear regression model of X and Y, as compared to a model that includes only previous observations of X. Since its introduction, G-causality has found wide application in economics and many other fields, including neuroscience and climatology [16, 38]. It is important to recognize that G-causality is a statistical formulation of causality, such that a significant G-causality interaction does not by itself imply the presence of a corresponding physical interaction [20].

To illustrate G-causality, suppose that the temporal dynamics of two time series $X_1(t)$ and $X_2(t)$ (both of length T) can be described by a bivariate autoregressive model:

$$\begin{aligned} X_1(t) &= \sum_{j=1}^{p} A_{11,j} X_1(t-j) + \sum_{j=1}^{p} A_{12,j} X_2(t-j) + \xi_1(t), \\ X_2(t) &= \sum_{j=1}^{p} A_{21,j} X_1(t-j) + \sum_{j=1}^{p} A_{22,j} X_2(t-j) + \xi_2(t), \end{aligned} \tag{1}$$

where p is the maximum number of lagged observations included in the model (the model order $p < T$), A contains the coefficients of the model, and ξ_1, ξ_2 are the residuals (prediction errors) for each time series. If the variance of ξ_1 (or ξ_2) is reduced by the inclusion of the X_2 (or X_1) terms in the first (or second) equation, then it is said that X_2 (or X_1) *G-causes* X_1 (or X_2). Assuming that X_1 and X_2 satisfy covariance stationarity (i.e., unchanging mean and variance), the magnitude of this interaction can be measured by the log ratio of the prediction error variances for the restricted (R) and unrestricted (U) models:

$$gc_{2\to 1} = \log \frac{\mathrm{var}(\xi_{1R(12)})}{\mathrm{var}(\xi_{1U})}, \tag{2}$$

where $\xi_{1R(12)}$ is derived from the model omitting the $A_{12,j}$ (for all j) coefficients in the first equation, and ξ_{1U} is derived from the full model. Importantly, G-causality is easy to generalize to the multivariate

(conditional) case in which the G-causality of X_1 is tested in the context of multiple variables $X_2,\ldots, X_N$ ($X_i \neq X_j$ for all $X_{i,j}$). In this case, X_2 G-causes X_1 if knowing X_2 reduces the variance in X_1's prediction error when the activities of all other variables $X_3,\ldots, X_n$ are also included in the regression model.

Extensions to standard G-causality include frequency-domain implementations, which are widely used in neurophysiology [18, 26], and a new *partial* G-causality, which exploits correlations between residuals in order to partly eliminate the influence of unmeasured variables [22]. In this article, however, I consider only the simple time-domain formulation given above.

2.2 G-autonomy

The framework of G-causality provides a straightforward means of operationalizing autonomy as self-determination, or self-causation. Instead of asking whether the prediction error of X_1 is reduced by including past observations of X_2, the G-autonomy measure asks whether the prediction error of X_1 is reduced by inclusion of its own past, given a set of external variables $X_{2,\ldots,N}$. That is, a variable X_1 is G-autonomous to the extent that its own past states help predict its future states over and above predictions based on past states of a set of external variables $X_{2,\ldots,N}$. Put simply, a variable is G-autonomous to the extent that it is dependent on its own history and that these dependencies are not accounted for by external factors.

Recalling Equation 1, if the variance of ξ_1 (or ξ_2) is reduced by the inclusion of the X_1 (or X_2) terms in the first (or second) equation, then it is said that X_1 (or X_2) is G-autonomous with respect to X_2 (or X_1). In other words, X_1 is G-autonomous if the coefficients in A_{11} are jointly significantly different from zero. As with G-causality, this can be tested by performing an F-test of the null hypothesis that $A_{11} = 0$, given assumptions of covariance stationarity on X_1 and X_2. By analogy with G-causality, the G-autonomy of X_1 with respect to X_2 is given by

$$ga_{X_1|X_2} = \log \frac{\mathrm{var}(\xi_{1R(11)})}{\mathrm{var}(\xi_{1U})}, \tag{3}$$

where $\xi_{1R(11)}$ is derived from the model omitting the coefficients $A_{11,j}$ (for all j) in the Granger equations.

2.3 G-emergence

G-emergence is a continuous version of weak emergence, in which a macro property is weakly emergent *to the extent that* it is not deducible from micro-level observations. G-causality and G-autonomy together provide the necessary ingredients for operationalizing G-emergence. According to G-emergence, a macro variable M is emergent from a set of micro variables $\mathbf{m}$ if and only if (i) M is G-autonomous with respect to $\mathbf{m}$ *and* (ii) M is G-caused by $\mathbf{m}$. A simple measure for the G-emergence of M from $\mathbf{m}$ is therefore given by

$$ge_{M|\mathbf{m}} = ga_{M|\mathbf{m}} \left(\frac{1}{N} \sum_{i=1}^{N} gc_{m_i \to M} \right). \tag{4}$$

This measure captures the basic intuition about weak emergence that it involves dependence on underlying processes, and that at the same time it involves autonomy from underlying processes [6]. Importantly, $ge_{M|\mathbf{m}}$ will be zero either if M is independent of $\mathbf{m}$ or if M is fully predicted by $\mathbf{m}$.

A criticism of G-emergence as measured using linear regression modeling is that a macro variable may appear to be G-emergent by virtue of being a nonlinear function of its micro-level components. Clearly, a satisfying measure of emergence should not rely on the failure of linear methods to detect nonlinear dependencies. Fortunately, it is easy to extend G-causality (and hence G-autonomy and G-emergence) to nonlinear situations. One simple method, given below, is to use a Taylor expansion. In the following

example the set of variables has been expanded to three in order to illustrate extension to the conditional ($n > 2$) case:

$$X_1(t) = \sum_{k=1}^{q}\sum_{j=1}^{p} A_{11,j,k}X_1^k(t-j) + \sum_{k=1}^{q}\sum_{j=1}^{p} A_{12,j,k}X_2^k(t-j) + \sum_{k=1}^{q}\sum_{j=1}^{p} A_{13,j,k}X_3^k(t-j) + \xi_1(t),$$

$$X_2(t) = \sum_{k=1}^{q}\sum_{j=1}^{p} A_{21,j,k}X_1^k(t-j) + \sum_{k=1}^{q}\sum_{j=1}^{p} A_{22,j,k}X_2^k(t-j) + \sum_{k=1}^{q}\sum_{j=1}^{p} A_{23,j,k}X_3^k(t-j) + \xi_2(t),$$

$$X_3(t) = \sum_{k=1}^{q}\sum_{j=1}^{p} A_{31,j,k}X_1^k(t-j) + \sum_{k=1}^{q}\sum_{j=1}^{p} A_{32,j,k}X_2^k(t-j) + \sum_{k=1}^{q}\sum_{j=1}^{p} A_{33,j,k}X_3^k(t-j) + \xi_3(t), \tag{5}$$

where q is the number of polynomial terms to be included in the Taylor expansion. Assuming that X_1 is a macro-level variable and $X_{2,3}$ are micro-level variables, the G-emergence of X_1 from $X_{2,3}$ is given by

$$ge_{X_1|X_2,X_3} = \log\frac{\mathrm{var}(\xi_{1R(11)})}{\mathrm{var}(\xi_{1U})} \times \frac{1}{2}\left(\log\frac{\mathrm{var}(\xi_{1R(12)})}{\mathrm{var}(\xi_{1U})} + \log\frac{\mathrm{var}(\xi_{1R(13)})}{\mathrm{var}(\xi_{1U})}\right), \tag{6}$$

where, following the previous convention, $\xi_{1R(ab)}$ is derived from the model omitting the A_{ab} coefficients in Equation 5. A value of the linear or nonlinear G-emergence can be considered statistically significant if the corresponding G-autonomy and G-causality measures are themselves statistically significant. This can be assessed by F-tests on the null hypothesis that the coefficients in A_{11} (G-autonomy) and $A_{12},\ldots, A_N$ (G-causality) are zero [18, 19].

The concept of G-emergence does not depend on using a particular method for nonlinear regression. There exist other, more sophisticated methods than Taylor expansions, which can be less sensitive to noisy observations and which involve fewer parameters. For example, Ancona and colleagues have shown that radial basis functions can serve as effective regression kernels for measuring nonlinear Granger causality [3]. However, for present purposes the Taylor method is preferable because it is simple to describe and to implement, statistical significance can easily be assessed, and it supplies an explicit formula for G-emergence (Equation 6).

3 Simulation Models: G-autonomy

3.1 Validation of G-autonomy

G-autonomy can be validated by application to time series that have different self-causation characteristics by construction. Consider the multivariate system X, illustrated in Figure 1:

$$\begin{aligned}
X_1(t) &= \xi_1(t),\\
X_2(t) &= aX_1(t-1) + (1-a)\xi_2(t),\\
X_3(t) &= bX_3(t-2) + (1-b)\xi_3(t),\\
X_4(t) &= cX_3(t-1) + (1-c)\xi_4(t),\\
X_5(t) &= dX_1(t-2) + dX_5(t-2) + d\xi_5(t).
\end{aligned} \tag{7}$$

In this system, $\xi_{1,\ldots,5}$ are independent white noise processes, the parameters a, b, c are all equal to 0.5, and $d = 1/3$. To calculate G-autonomy values, 10,000 values of each time series were generated and

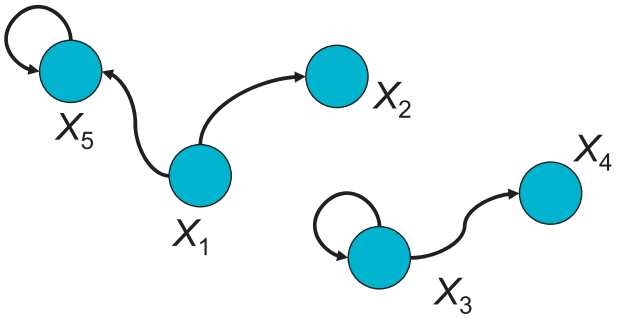

Figure 1. Schematic of constructed causal relations in the system X described in Equation 7.

an order-2 multivariate regression model was computed, using ordinary least squares [24]. Table 1 shows the G-autonomy values for each variable. For comparison, the G-causality values among variables are also shown.

Table 1 shows that G-autonomy satisfies several intuitive requirements for a quantitative measure of autonomy. X_1 has a G-autonomy of zero, which is expected because it is a white noise process. X_2 also has a G-autonomy of zero, because it is a combination of a white noise process and influences from another white noise process, X_1. X_3 has a high G-autonomy because its future evolution can be partially predicted from its own past in a way that does not depend on the other variables (X_1, X_2, X_4, X_5). X_4 has a low G-autonomy because, although its future can be partially predicted by its own past, the accuracy of this prediction is due entirely to influence from X_3. Finally, X_5 has a high G-autonomy because, like X_3, it contains an autoregressive component that is not accounted for by the other variables. Thus, in this example and in general, a variable has high G-autonomy if and only if its future can be better predicted from its own past after taking into account the predictive contribution of a set of external variables.

It is instructive to compare the G-autonomy values with the corresponding G-causality values. There are three significant G-causality interactions: $X_1 \rightarrow X_2$, $X_3 \rightarrow X_4$, and $X_1 \rightarrow X_5$, as expected from the construction of X (see Figure 1). Notice that while both X_1 and X_3 exert G-causal influences on other variables, X_3 is G-autonomous whereas X_1 is not. Also, X_1 G-causes both X_2 and X_5, but X_5 is G-autonomous whereas X_2 is not. Therefore, knowing the set of G-causalities among variables is not sufficient to know the G-autonomies of each variable with respect to others in the set.

3.2 G-autonomy in a Predator-Prey Model

As a more general test of the utility of G-autonomy, I now consider a simple agent-based model in which a *predator* agent moves in a 2D toroidal plane populated by mobile *prey*. This model is used to investigate

Table 1. G-autonomy and G-causality values for an illustrative data set generated from X (Equation 7). G-autonomies are along the diagonal (boldface). G-causalities are arranged so that the column variable G-causes the row variable. Note that G-causalities are *conditional* because they are derived from a fully multivariate model. Statistically significant values ($p < 0.01$) are indicated with an asterisk.

	X_1	X_2	X_3	X_4	X_5
X_1	**0**	0.03	0	0.5	0
X_2	8.4*	**0**	0	0	0
X_3	0.6	0.4	**7.9***	0.3	0
X_4	0	0	7.7*	**0**	0
X_5	7.3*	0.7	0.3	0	**7.0***

whether G-autonomy applied to predator behavior captures intuitive characteristics of autonomous behavior as self-determination, and whether evolutionary adaptation influences G-autonomy in a predator-prey context.

3.2.1 Experimental Conditions

In the model there are two types of prey, red and green, and the predator has two internal batteries, one corresponding to each prey type. Four experimental conditions are compared:

1. Predator moves randomly.
2. Predator behavior is determined by environment.
3. Predator behavior is determined by a combination of environment and internal state.
4. Predator behavior is controlled by an evolved neural network.

In each condition there are three red prey and three green prey. Prey move at a speed of 3 distance units per time step (u/ts). The environment is 200 u by 200 u in size. In all conditions prey change heading each time step by an angle randomly chosen from the range $[-\pi/8, \pi/8]$. If the predator comes within 10 u of a prey, the corresponding battery is fully replenished. Batteries otherwise deplete by 1 u/ts from an initial value of 300.

In all conditions the predator moves at a speed of 4 u/ts, that is, slightly faster than prey. In condition 1, the predator changes direction each time step by an angle randomly chosen from the range $[-\pi/8, \pi/8]$. In condition 2, the predator adjusts its heading at each time step to point toward the red prey that was closest at the beginning of the trial. In condition 3, the predator adjusts its heading each time step to point toward the nearest red prey if the red battery level is lower than the green battery level, or otherwise toward the nearest green prey. In condition 4, the predator heading is controlled by a feedforward neural network optimized using a genetic algorithm to maximize the average level of both batteries, as described in the Appendix.

3.2.2 Results and Analysis

Each condition of the model was run for 10 trials of 10,000 ts each (following evolution in condition 4). From each trial, seven time series were generated, corresponding to the trajectories of the predator and the six prey. To prevent edge effects, each time series ($a_1, \ldots, a_7$) consisted of the average of the horizontal and vertical displacement of the predator or prey from the midlines of the environment, that is,

$$a_i(t) = 0.5(|x_i(t) - 100| + |y_i(t) - 100|).$$

Each time series was then first-order differenced [i.e., $a(t) \rightarrow a(t) - a(t-1)$] in order to ensure covariance stationarity [36]. After differencing, all time series from all conditions were covariance stationary ($p < 0.01$, Dickey-Fuller test). Each resulting data set was used to construct a multivariate autoregressive model of order $p = 4$. This model order was chosen according to the Akaike information criterion (AIC) [1] (the mean minimum AIC, computed from all data sets, was 3.9).[2] To verify that the order-4 models sufficiently described the data, it was noted that the mean adjusted residual-sum-square RSS_{adj} was sufficiently high (0.86 ± 0.07). Each model was then used to calculate predator G-autonomy values (Equation 3). G-autonomies were averaged across the 10 trials in each condition. For comparison, a nonlinear G-autonomy was also calculated via a Taylor series expansion method as described in Section 2.3.

2 To check robustness, the analysis was repeated with $p = 8$, because the maximum lowest AIC (among conditions) was 7.8. The results were unchanged from $p = 4$.

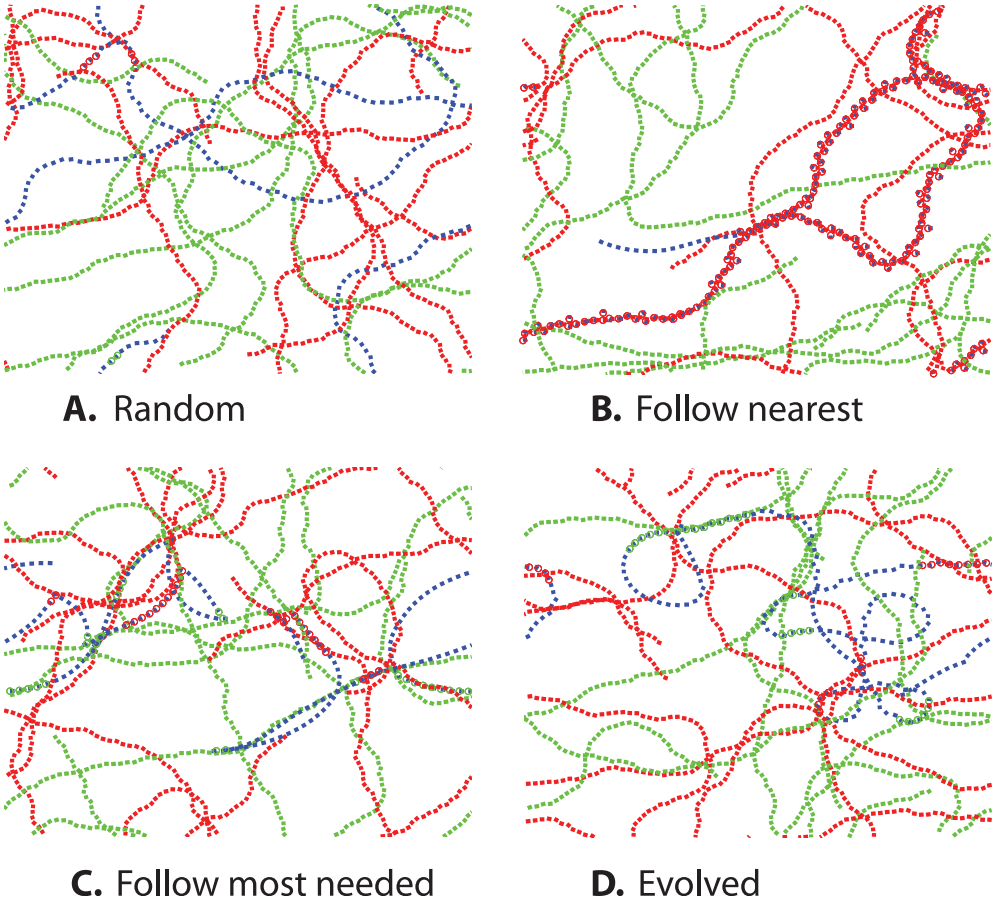

Figure 2. Example predator and prey trajectories from each condition (200 ts are shown). A, predator moves randomly; B, predator behavior is determined by environment; C, predator behavior is determined by a combination of environment and internal state; D, predatory behavior is controlled by an evolved neural network. Blue shows predator trajectory; red and green show red and green prey trajectories respectively. (Color online.)

Figure 2 shows sample trajectories of predator and prey from each condition. All conditions exhibit rich behavioral dynamics. By inspection, only condition 2 (follow the nearest red prey) appears clearly distinguishable from the others, as might be expected given the simplicity of the rule governing predator behavior in this condition.

Figure 3 shows the average predator G-autonomy values in each condition, for both linear (A) and nonlinear (B) analyses. Random predator movement (condition 1) produces high G-autonomy, as expected, because random movement in this model is not a white noise process; rather, it implies that the predator changes direction unpredictably, so that the best predictors of future predator position are recent past positions.[3] Also as expected, condition 2 produces low G-autonomy, because the predator position is now well predicted by past positions of a particular prey.

Predator G-autonomy in condition 3 is higher than in condition 2, but lower than the upper bound provided by random movement (condition 1). This reflects the fact that predator behavior in this condition is driven both by the position of a nearby prey and by the relative levels of the two internal batteries. Notably, G-autonomy in condition 4 is significantly higher than in condition 3, suggesting that evolutionary adaptation in the model leads to increased G-autonomy. A possible

3 Note that $a_{\text{predator}}(t)$ in this condition is not a true random walk. Therefore, even after first-order differencing, past positions remain useful predictors of present positions.

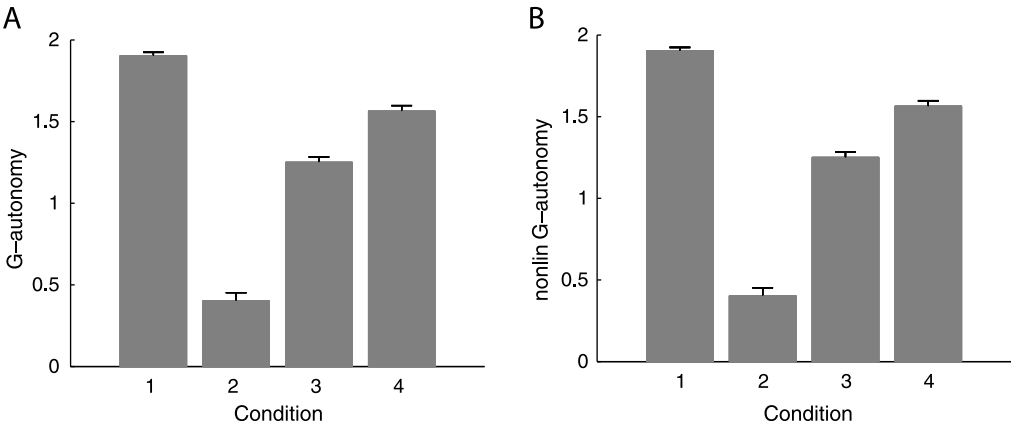

Figure 3. A: Mean predator G-autonomy for each condition. Error bars show standard deviations. All comparisons are statistically significant (two-tailed *t*-test, $p < 0.001$, after Bonferroni correction for multiple comparisons). B: Mean predator nonlinear G-autonomy for each condition.

explanation for this is that the evolved controller has the opportunity to take the internal state into account at every time step, and not only at those time steps at which the lowest battery level changes from red to green or vice versa as is the case in condition 3. G-autonomy in condition 4 remains lower than in condition 1 because predator position can still be partially predicted by prey position. Results from linear and nonlinear analyses were qualitatively identical.

3.3 G-Autonomy Summary

The G-autonomy measure quantifies and amplifies the concept of autonomy as self-determination. The measure behaves appropriately both in a constructed system and in a simulation model of predation. A striking finding in the latter case was that evolutionary adaptation enhanced predator G-autonomy. This is consistent with the general notion that evolutionary processes promote autonomy [30, 34] and opens the way toward explicit modeling of the relation between evolution and autonomy.

4 Simulation Models: G-Emergence

A canonical example of emergence is a flock of starlings wheeling in the sky before they roost: The flock seems to have a shape and trajectory of its own, which transcends those of the individual birds (Figure 4).

Figure 4. A flock of starlings over Brighton pier. Photo courtesy of Eduardo Izquierdo.

In a seminal ALife work, Reynolds [33] showed that visually compelling bird flocking can be simulated by combining three simple rules for simulated birds (boids):

- *Aggregation.* Each boid tends to fly toward the perceived center of mass (CM) of the flock.
- *Avoidance.* Each boid tends to avoid colliding with other nearby boids.
- *Matching.* Each boid tends to align its velocity with that of other nearby boids.

Here, a simple boids simulation is used to test whether visually compelling flocking correlates with high G-emergence of the CM of the flock (the macro variable) with respect to the trajectories of the individual boids (the micro variables).

$N = 10$ boids are simulated in a toroidal square environment of length 200 u. Boids are initialized with positions and velocities randomly chosen from the range [0, 200] (x, y position), [0, 2π] (heading), and [3, 9] (speed). At each time step the heading α_i and speed s_i of each boid i are updated synchronously according to

$$\alpha_i = \alpha_i + a_1\theta_1 + a_2(\pi + \theta_2) + a_3\theta_3 + r_1,$$
$$s_i = s_i + a_4 ds + r_2,$$

where θ_1 is the bearing to the perceived CM (i.e., the CM not including boid i), θ_2 is the bearing to the nearest boid, θ_3 is the bearing to the mean heading of all other boids within a 20-u range, $\overline{ds}$ is the difference between the speed of boid i and the mean speed of all other boids within 20 u, and r_1 and r_2 are random numbers in the range $[-0.01, 0.01]$. The parameter vector a (all $a \in [0, 1]$) determines the relative contribution of each factor. Toroidal distances are calculated in the standard way, as the minimum distance either across, or not across, the boundary. CM positions are calculated iteratively in order to minimize the summed toroidal distances to the boids (i.e., not as the average boid position, which can lead to boundary artifacts).

Three different conditions were tested. Condition R (random) produces near-random boid behavior ($a_R = [0.01, 0.01, 0.01, 0.01]$). Condition L (low) evokes poor flocking behavior by imposing a strong dependence on velocity matching; boids in this condition tend to move in semi-rigid formations ($a_L = [0.1, 0.1, 0.6, 0.6]$). Condition H evokes compelling flocking behavior; the parameter set ($a_H = [0.1, 0.3, 0.3, 0.3]$) was selected by hand. Examples of boid and CM trajectories from each condition are shown in Figure 5. Although static images do not fully capture the dynamic nature of flocking, it is clear that boid trajectories in condition H are more flocklike than those in conditions L and R.

4.1 Results and Analysis

For each condition the boid simulation was run 25 times with each run lasting 5000 ts; for each run the x, y coordinates of each boid and the global CM were recorded. Several preprocessing steps were carried out prior to calculation of G-emergence. As in the predator-prey model, each x, y coordinate pair was transformed into a single variable reflecting distance from the center of the environment. The first 500 data points were removed to eliminate initial transients, and each resulting time series was transformed into its zero-mean equivalent. Finally, each time series was first-order differenced in order to ensure covariance stationarity. Following preprocessing, for each run in each condition both linear and nonlinear G-emergence of the CM were computed using a model order $p = 5$ and (for the nonlinear analysis) a polynomial order $q = 3$. The model order was selected based on the average AIC across all 75 runs.

Figure 5 (top left) shows the mean linear and nonlinear G-emergence of the CM in each condition. Confirming the prediction that high-G-emergence tracks compel flocking, both linear and nonlinear measures show significantly higher values of G-emergence in condition H than in conditions L and R. All values of G-emergence in conditions H and L were significant ($P < 10^{-5}$ for G-autonomy and G-causality, two-tailed t-test); those in condition R were not.

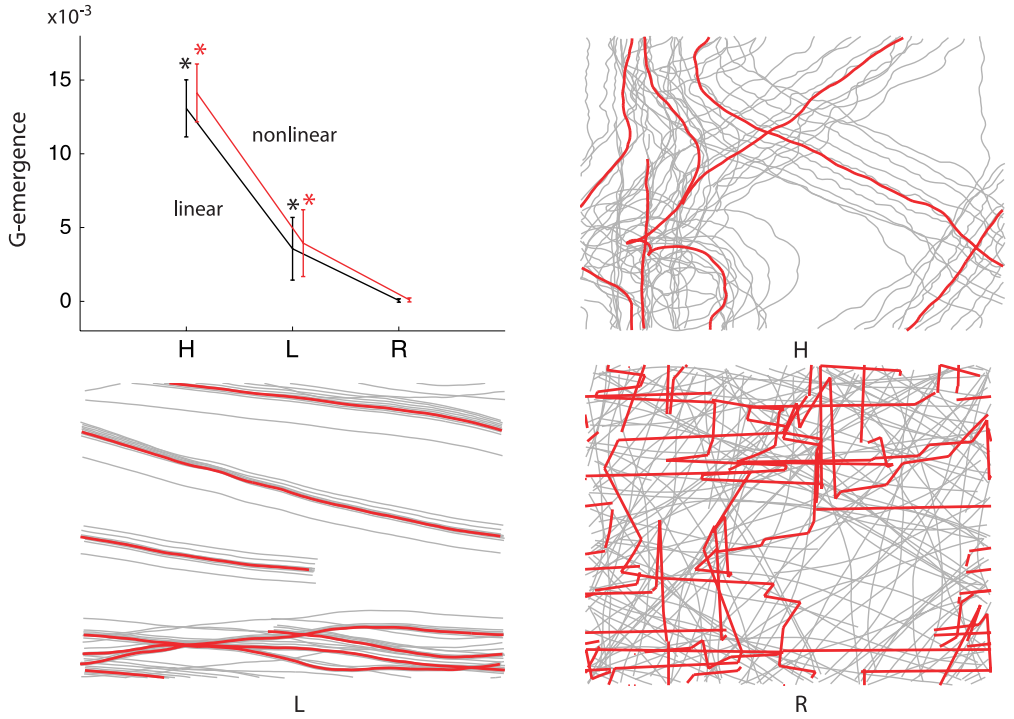

Figure 5. G-emergence of the center of mass (CM) of a boid flock. Top left: Mean and standard deviation of linear and nonlinear G-emergence by condition (asterisks show statistical significance). Other panels: Example trajectories (500-ts segment) of the boids (gray) and CM (red) in conditions H (high G-emergence), L (low G-emergence), and R (random). (Color online.)

To test the behavior of G-emergence across different parameter combinations in the boids model, linear and nonlinear G-emergence values were calculated for each parameter vector in the space $a_{(1,2,3)} \in \{0.0, 0.1, \ldots, 1.0\}$. Parameters a_3 and a_4 were yoked together because they both influence the same rule (velocity matching), and three evaluations were carried out for each vector, requiring a total of $11 \times 11 \times 11 \times 3 = 3993$ evaluations. Figure 6 shows G-emergence for three orthogonal cross sections through the three-dimensional parameter space; in each cross section the vector corresponding to a_H (condition H) is marked by the intersection of the green lines.

Several aspects of the above cross sections are notable. First, as with the G-autonomy results, linear and nonlinear G-emergence are strongly correlated, suggesting that even linear measures can provide insight into emergent properties in some complex systems. Second, in most regions of parameter space G-emergence changes smoothly, suggesting it is a robust measure. However, some regions show sharp transitions, for example between some vectors with $a_1 = 0$ and neighboring vectors. The sensitivity of G-emergence to these transitions indicates that it can usefully identify parameter regions of complex models in which nontrivial weak emergence is present.

4.2 Downward Causation

A common intuition regarding emergence is that it involves *downward* causation from macro levels to micro levels. For proponents of strong emergence, downward causation is in fact an essential aspect of what it means to be emergent [27]. However, physical interpretations of downward causation pose tricky ontological problems, for example, how to resolve competing micro and macro causes [7]. G-emergence, being statistically defined, provides an ontologically innocent alternative according to which downward causation is reflected by G-causality from the macro variable(s) to the micro variable(s).

Figure 7 shows downward (Granger) causation from the global CM to the individual boid trajectories, for both linear and nonlinear G-causality measures. Averages are taken across all boids and across all 25 runs in each condition. Consistent with an association between emergence and downward causation, both measures of downward causation are significantly higher in condition H than in condition R or L.

The above result aside, it seems possible in principle for weak emergence to occur without downward causation (strong emergence requires downward causation by definition). Having separately applicable measures of weak emergence and downward causation makes it possible to explore conditions in which emergence and downward causation do not occur together, potentially refining and deepening the concept of emergence.

4.3 Emergence Summary

G-emergence is a quantitative, intuitive, and practically straightforward measure of weak emergence. It is based on the intuition that emergent properties are both *dependent on* and *autonomous from* their components [6, 8] and is operationalized using linear and nonlinear time series analysis. In a simulation of bird flocking, visually compelling flocking behavior is accompanied by high G-emergence as compared

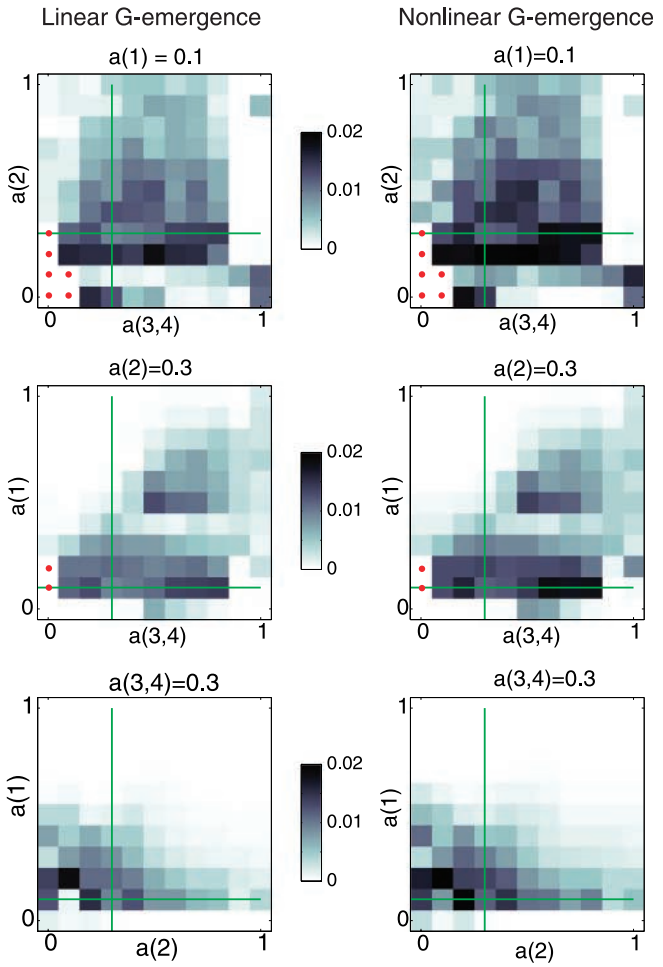

Figure 6. Parameter space of the boids model. The parameter vector a_H is indicated by the intersection of the green lines. Gray scale shows average linear and nonlinear G-emergence of the global CM. Each value is the average of three evaluations of 5000 ts each. Red dots indicate parameter combinations that lead to reliably nonstationary time series. (Color online.)

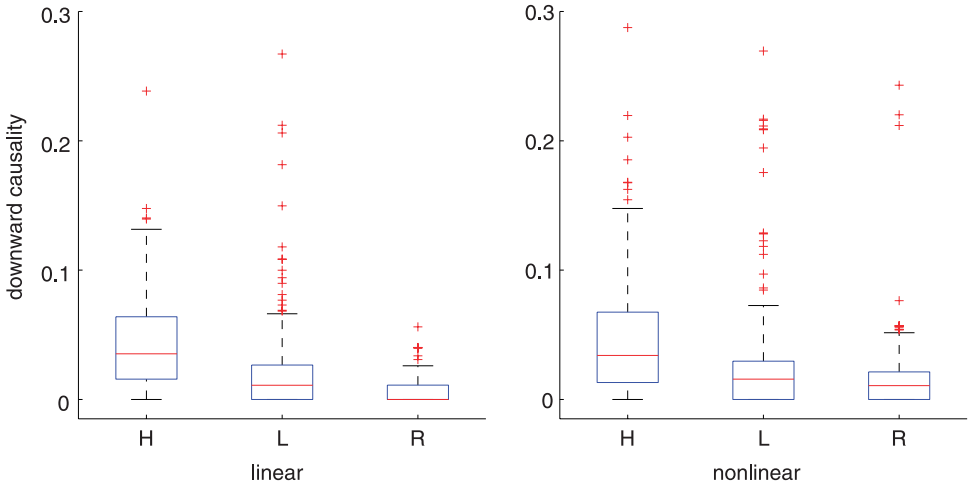

Figure 7. Downward causation is significantly higher in condition H than in condition L or R. Boxplots show linear and nonlinear G-causality from the global CM to individual boids, calculated separately for each boid for all 25 runs in each condition (i.e., 250 values per boxplot). Non-significant causalities were set to zero (nominal threshold of 0.01, Bonferroni corrected to 10^{-5}). Resulting distributions are non-normal, and differences between conditions were tested using the Wilcoxon rank sum test. For both linear and nonlinear analyses all pairwise comparisons among medians were significantly different ($p < 10^{-3}$). Each boxplot shows lower quartile, median, and upper quartile values; whiskers show range of remaining data, and + denotes outliers.

to random movement or flight in rigid formations. High G-emergence is also accompanied by downward (Granger) causation from the flock to each boid, though this may not be the case for all systems.

5 Discussion

Having the ability to measure a phenomenon is an essential step toward its effective scientific description [13]. Moreover, measures do not exist in a vacuum; they must prove themselves by iteratively building knowledge in the context of theoretical frameworks. The ultimate virtue in a measure is not its a priori robustness, but its ability to build on intuitions, identify interesting divides in nature, and then correct the foundations on which it was built [43]. In this article I have introduced and illustrated measures of two important properties of complex systems—autonomy and emergence—in precisely this spirit. A MATLAB (Natick, MA) toolbox containing functions for calculating G-causality and G-autonomy (and hence also G-emergence) is freely available from the author's web site (http://www.anilseth.com); the toolbox is described elsewhere [41].

5.1 Autonomy

G-autonomy provides an intuitive and easily applicable quantitative measure of autonomy based on the concept of self-determination. It can be applied not only to behavioral data, but also to time series data reflecting internal variables (e.g., neuronal activity, metabolic activity), or indeed to any multivariate data set in which autonomy is of interest. A G-autonomous variable is not simply one with a significant autoregressive component; rather, it is a variable with an autoregressive component that remains significant after taking into account the predictive ability of a set of external variables; G-autonomy therefore provides additional information to that found in a set of G-causalities among variables.

An important open question is the relation between G-autonomy and autonomy as conceived within the autopoietic framework, that is, autonomy as "organizational closure" or "the condition of subordinating all changes to the maintenance of the organization" [29, p. 135]. The two concepts seem closely related. For example, from an evolutionary perspective, selection may favor both behavioral autonomy and organizational closure, with the latter as a precursor to the former. Another point of connection is that organizationally closed descriptions may be emergent with respect to lower descriptive levels, for the reason

that organizational closure implies some indifference toward internal "microscopic" operations [2]. Further work is required to examine these possibilities.

Another important open question is that of autonomy in reciprocally causal systems. What does it mean for an agent to be autonomous if its behavior is both caused by and causal to an external environment? Finally, measurement of G-autonomy requires that all elements of a system show changing activity over time. It is not possible at present to measure the G-autonomy of an agent with respect to a static feature of the environment.

5.1.1 Relation to Other Measures

A similar approach, though not one tested in agent-based models, has been developed by Bertschinger and colleagues [9]. These authors propose measuring the conditional mutual information between consecutive states of a system, conditioned on the history of the environment. Information-theoretic approaches are naturally nonlinear and therefore are in principle more general than autoregressive frameworks. For example, *transfer entropy* [35] may be considered to be a more general version of Granger causality. However, information-theoretic measures require adequate sampling of potential outcomes in order to estimate probability distributions. Such sampling can be very challenging for multivariate systems (>2 variables). Scaling to multivariate systems is in general less severe for autoregressive approaches, though it can still be challenging [16]. Furthermore, autoregressive approaches benefit from standard methods for testing for statistical significance and have natural spectral interpretations (see Section 2.1).

Another difference between the two approaches is that Bertschinger et al. propose two distinct measures, one (as stated above) for cases in which the system cannot influence its environment, and another (the non-conditioned mutual information between consecutive system states) for cases in which the system has full control over the environment. G-autonomy is applicable equally in both cases, without modification, as well as in the common intermediate case in which the system and the environment mutually constrain and influence each other.

5.2 Emergence

G-emergence operationalizes a continuous version of weak emergence, such that a macro property is weakly emergent *to the extent that* it is both causally dependent on and autonomous from its micro-level components. Because G-emergence is based on a statistical interpretation of causality, it sidesteps conceptual pitfalls such as competition among micro and macro causes, while providing an objective and graded assessment of the nontriviality of micro-to-macro inferential pathways. In contrast to Bedau's concept of "incompressibility of explanation" G-emergence does not require proving a negative; moreover, from the perspective of measurement a continuous value is more useful than a binary, or categorical, classification.

Under what circumstances could G-emergence be high? A macro variable could be G-emergent from a set of micro variables if there are hidden, or latent, influences, that is, relevant micro causal factors that are not represented in the regression model. However, even if all micro causal factors are represented, G-emergence can still exceed zero because of dependence on the prediction algorithm used. It is plausible, and indeed necessary for G-emergence to be useful in practice, that macro variables vary in their epistemic transparency to a given prediction algorithm, relative to collections of corresponding micro variables. Clearly, if the system itself were used as the prediction "algorithm," then G-emergence would always be zero, since on the present view it is assumed that micro causes are jointly sufficient to give rise to macro properties.

5.2.1 Property Emergence versus Temporal Emergence

Intuitively, emergence refers either to a macro-level property that is "more than the sum of" the micro-level parts (*property*, or "synchronic," emergence) or to the appearance of a qualitatively distinctive new phenomenon over time (*temporal*, or "diachronic," emergence). Temporal emergence is well illustrated by the appearance of new morphological features during embryogenesis and development. Because a

temporally emergent process is by definition statistically nonstationary, it cannot be measured by G-emergence. Nonetheless, it is plausible that such processes are bracketed by statistically stationary periods with different G-emergence properties. In this way, G-emergence could be used indirectly to infer temporal emergence.

5.2.2 Phase Transitions

Physicists have recently become interested in the onset of collective behaviors among boidlike self-propelling particles [21, 48]. In such systems, phase transitions can be observed among *gaseous* phases (each particle moves independently), *liquid* phases (particles move collectively but still diffuse with respect to each other), and *solid* phases (particles move collectively and remain fixed with respect to each other). Plausibly, these phases correspond respectively to conditions R, H, and L of the model described in Section 4, and the sharp boundaries noted in Figure 6 may correspond to phase transitions. However, phase transition analyses tend to focus on the dynamics of transition and assume that emergent behavior is phenomenologically obvious in some phases and is absent in others. In contrast, G-emergence focuses on detecting the *degree* of emergence by making physical measurements on a system.

5.2.3 Relation to Other Measures

The intuition that variations in predictive ability may be important in defining macro-level properties is shared by Shalizi and Moore [44]. However, these authors focus on clarifying the concept of a macro state, and they do not explicitly combine measures of autonomy and causal dependence. Rather, one process is called emergent from another if it has a higher *predictive efficiency* than the process it derives from. Their measure of predictive efficiency is based on information-theoretic model reconstruction (the epsilon-machine concept [14]), which is powerful but less easy to apply in practice than the G-causality approach described here (see also the related concept of *statistical complexity* [15]). A related approach is taken by Polani [31], in which an "emergent description" involves a further step of decomposing systems into independent informational subcomponents.

According to the *contextual emergence* of Atmanspacher, derivation of macro-level properties requires knowledge of micro-level properties and of contingent contextual conditions, the latter defined at the macro level and implemented in terms of stability criteria according to a dynamical systems analysis [4]. This concept diverges from the doctrine of causal fundamentalism (or supervenience) by proposing that micro-level properties offer necessary but not sufficient conditions for deriving macro-level properties. As with weak emergence, contextual emergence addresses questions of epistemology rather than ontology [4]. By contrast, Bar-Yam offers an explicit measure of strong (ontological) emergence, which is based on measuring the entropy of a system at multiple scales [5]. Oscillations in "multiscale variety" are suggested to reveal constraints on the values of multiple variables that are not present among subsets of these variables, and the existence of such constraints is taken to indicate strong emergence.

5.3 Strong Emergence and Consciousness

The philosopher David Chalmers has made explicit a recurring idea, which is that there is exactly one clear case of a strongly emergent phenomenon, and that is the phenomenon of consciousness [12]. It seems that two commonly held intuitions about consciousness drive this suspicion. First, the idea that even complete knowledge of the physical interactions sustained by brains will not provide an understanding of what it is like to have a conscious experience: This is the infamous "hard problem" of consciousness [12]. Second, the intuition that conscious states have causal efficacy in the world [32], as exemplified by the notion of free will, but which runs through all aspects of consciousness; after all, why have experiences at all if they do not do anything? These intuitions map cleanly onto the defining features of strong emergence, namely, that macro-level properties in principle cannot be identified from micro-level observations, and that macro-level properties have irreducible causal powers.

These intuitions can, however, be challenged. First, to expect a scientific resolution to the "hard problem" as it is presently conceived may be to misunderstand the role of science in explaining nature. A scientific theory cannot presume to replicate the experience it describes or explains; a theory of a

hurricane is not a hurricane [42]. If the phenomenal aspect of experience is irreducible, so is the fact that physics has not explained why there is something rather than nothing, and this has not prevented physicists from laying bare many mysteries. Second, consciousness can be functionally efficacious without assuming downward causation. It is entirely plausible that certain neural mechanisms support useful functions by virtue of the fact that they entail conscious experiences [40]. For example, the neural mechanisms underlying consciousness may serve to integrate large amounts of information over short time periods, leading to functionally effective high-dimensional discriminations among a large repertoire of sensorimotor scenes [45]. Such information integration may entail consciousness in just the same way that the molecular structure of hemoglobin entails a particular spectroscopic response: It simply could not be otherwise [17]. Moreover, experiences of free will and volition are just experiences like any other, and there is a wealth of experimental evidence showing, unsurprisingly, that awareness of a voluntary action is preceded by recognizable signatures in neural activity [23, 28]. Together, these points suggest that the association of consciousness with strong emergence does not rest on solid ground.

In contrast, it is very likely that the connection between neural mechanism and conscious experience involves *weak* emergence in many ways. A striking feature of conscious experience is that it seems more than the sum of its parts (each conscious experience is a unity) and that it has a vivid temporality (William James' "stream of consciousness"). Models of consciousness that can be analyzed in terms of weak emergence therefore have the potential to *explain* features of phenomenology in terms of dynamic processes at the level of neural mechanism. The development and experimental testing of such *explanatory correlates* [39] is a highly promising avenue toward a scientific description of consciousness. For example, one could hypothesize that the extent to which a conscious experience includes a volitional component will correlate with the measurable downward G-causality from macro-level descriptions of brain dynamics relevant to consciousness to corresponding micro-level descriptions. It is exciting to consider that measures of weak emergence may eventually find utility in accounting for apparent free will and in crossing the explanatory gap between neural mechanism and phenomenal experience.

Acknowledgments

The author is supported by EPSRC leadership fellowship EP/G007543/1. Many thanks to Eduardo Izquierdo for the starling photo. This article is dedicated to Sir Clive Granger, who passed away on May 27, 2009.

References

1. Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.
2. Allefeld, C., Atmanspacher, H., & Wackermann, J. (2009). Mental states as macrostates emerging from brain electrical dynamics. *Chaos*, *19*, 015102.
3. Ancona, N., Marinazzo, D., & Stramaglia, S. (2004). Radial basis function approaches to nonlinear Granger causality of time series. *Physical Review E*, *70*, 056221.
4. Atmanspacher, H., & beim Graben, P. (2009). Contextual emergence. *Scholarpedia*, *4*, 7997.
5. Bar-Yam, Y. (2004). A mathematical theory of strong emergence using multiscale variety. *Complexity*, *9*, 15–24.
6. Bedau, M. (1997). Weak emergence. *Philosophical Perspectives*, *11*, 375–399.
7. Bedau, M. (2003). Downward causation and the autonomy of weak emergence. *Principia*, *6*, 5–50.
8. Bedau, M. (2008). Is weak emergence just in the mind? *Minds and Machines*, *18*, 443–459.
9. Bertschinger, N., Olbrich, E., Ay, N., & Jost, J. (2008). Autonomy: An information theoretic perspective. *Biosystems*, *91*, 331–345.
10. Boden, M. (1996). Autonomy and artificiality. In M. Boden (Ed.), *The philosophy of artificial life* (pp. 95–108). Oxford, UK: Oxford University Press.
11. Broad, C. (Ed.) (1925). *The mind and its place in nature*. Oxford, UK: Routledge and Kegan Paul.

12. Chalmers, D. (2006). Strong and weak emergence. In P. Clayton & P. Davies (Eds.), *The re-emergence of emergence*. Oxford, UK: Oxford University Press.

13. Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford, UK: Oxford University Press.

14. Crutchfield, J. (1994). The calculi of emergence: Computation, dynamics, and induction. *Physica D*, *75*, 11–54.

15. Crutchfield, J. P., & Young, K. (1989). Inferring statistical complexity. *Physical Review Letters*, *63*, 105–108.

16. Ding, M., Chen, Y., & Bressler, S. (2006). Granger causality: Basic theory and application to neuroscience. In S. Schelter, M. Winterhalder, & J. Timmer (Eds.), *Handbook of time series analysis* (pp. 438–460). New York: Wiley.

17. Edelman, G. (2003). Naturalizing consciousness: A theoretical framework. *Proceedings of the National Academy of Sciences of the U.S.A.*, *100*, 5520–5524.

18. Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, *77*, 304–313.

19. Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424–438.

20. Granger, C. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control*, *2*, 329–352.

21. Gregoire, G., Chate, H., & Tu, Y. (2003). Moving and staying together without a leader. *Physica D*, *181*, 151–170.

22. Guo, S., Seth, A., Kendrick, K., Zhou, C., & Feng, J. (2008). Partial Granger causality: Eliminating exogenous inputs and latent variables. *Journal of Neuroscience Methods*, *172*, 79–93.

23. Haggard, P. (2008). Human volition: Towards a neuroscience of free will. *Nature Reviews Neuroscience*, *9*, 934–946.

24. Hamilton, J. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

25. Jackson, F., & Pettit, P. (1992). In defence of explanatory ecumenism. *Economics and Philosophy*, *8*, 1–21.

26. Kaminski, M., Ding, M., Truccolo, W., & Bressler, S. (2001). Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics*, *85*, 145–157.

27. Kim, J. (1999). Making sense of emergence. *Philosophical Studies*, *95*, 3–36.

28. Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*, 529–566.

29. Maturana, H., & Varela, F. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht, The Netherlands: D. Reidel.

30. Moreno, A., Etxeberria, A., & Umerez, J. (2008). The autonomy of biological individuals and artificial models. *Biosystems*, *91*, 309–319.

31. Polani, D. (2006). Emergence, intrinsic structure of information, and agenthood. *Interjournal Complex Systems*, 1937.

32. Popper, K., & Eccles, J. (1977). *The self and its brain*. Berlin: Springer.

33. Reynolds, C. (1987). Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics*, *21*, 25–34.

34. Rosslenbroich, B. (2006). The notion of progress in evolutionary biology. The unresolved problem and an empirical suggestion. *Biology and Philosophy*, *21*, 41–70.

35. Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, *85*, 461–464.

36. Seth, A. (2005). Causal connectivity of evolved neural networks during behavior. *Network: Computation in Neural Systems*, *16*, 35–54.

37. Seth, A. (2007). Granger causality. *Scholarpedia*, *2*, 1667.

38. Seth, A. (2008). Causal networks in simulated neural systems. *Cognitive Neurodynamics*, *2*, 49–64.

39. Seth, A. (2009). Explanatory correlates of consciousness: Theoretical and computational challenges. *Cognitive Computation*, *1*, 50–63.

40. Seth, A. (2009). Functions of consciousness. In W. Banks (Ed.), *Springer Encyclopedia of Consciousness, Volume 1* (pp. 279–293). Berlin: Springer Verlag.

41. Seth, A. K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *Journal of Neuroscience Methods*, *186*, 262–273.

42. Seth, A. K., & Edelman, G. (2009). Consciousness and complexity. In B. Meyer (Ed.), *Springer Encyclopedia of Complexity and Systems Science*, Vol. 2 (pp. 1424–1443). Berlin: Springer Verlag.

43. Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*, 314–321.

44. Shalizi, C., & Moore, C. (2006). What is a macrostate: Subjective observations and objective dynamics. http://arxiv.org/abs/condmat/0303625.

45. Tononi, G., & Edelman, G. (1998). Consciousness and complexity. *Science*, *282*, 1846–1851.

46. van Gulick, R. (2001). Reduction, emergence and other recent options on the mind-body problem. *Journal of Consciousness Studies*, *8*, 1–34.

47. Varela, F. (1979). *Principles of biological autonomy.* Burlington, MA: Elsevier.

48. Vicsek, T., Czirok, A., Ben-Jacob, E., Cohen, I., & Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, *75*, 1226.

Appendix: Predator Neural Network and Genetic Algorithm

In condition 4 of the predator-prey model (Section 3), predators are simulated robots with six sensors, four of which are grouped in two left-right pairs. Two sensor pairs respond to the nearest red and green prey, and the response of each pair decays linearly with distance. If the nearest prey is to the left (right) of the predator, the first (second) sensor of the corresponding pair increases its response by a factor $0.5(|\theta|/(\pi/2))$, where θ is the bearing to the prey (range $[-\pi, \pi]$). If $|\theta| > \pi/2$, the corresponding sensor pair does not respond. The third sensor pair responds linearly to the level of red and green batteries. All sensor activities are linearly scaled to range from -3 to $+3$.

Predators are controlled by an 11-node feedforward neural network. The six sensors determine the input to six input neurons. The input neurons are fully connected to a layer of three intermediate neurons, which are fully connected to the two output neurons. Each neuron implements a sigmoidal transfer function mapping an input range $[-10, 10]$ onto an output of range $[0, 1]$. The motor output neurons are scaled to range from 0 to 5 to set the left and right wheel speeds. The forward speed is fixed at 4 u/ts to enable comparison with other experimental conditions, and the angular velocity is calculated as the difference between the wheel speeds divided by the robot diameter (5 u).

Network parameters are encoded as real numbers (range $[0, 1]$) on a 29-element genotype. The first five alleles specify biases for the intermediate neurons and output neurons (range $[-3, 3]$). The remaining 23 elements specify connection weights (range $[-3, 3]$). The fitness of each genotype is calculated as the average of five evaluations. The fitness of each evaluation is calculated as the summed levels of the two batteries for the lifetime of the predator. If either battery reaches zero, the evaluation is terminated (this constraint is removed during data analysis, for which it is important to obtain time series of consistent and extended length). Otherwise, evaluations last for 1000 ts. Positions and headings of predator and prey are initialized randomly at the start of each evaluation.

A genetic algorithm was used to evolve a population of 30 genotypes for 250 generations. The algorithm used stochastic rank-based selection with elitism and with a mutation probability of 0.09 per allele. Each mutation changed the allele value by a number randomly chosen from the range $[-0.2, 0.2]$ with reflection at the boundaries (0,1). High fitness was reliably achieved after approximately 100 generations.