Creative leaps in musical ecosystems: early warning signals of critical transitions in professional jazz

Matt Setzler*[1] msetzler@indiana.edu
Tyler Marghetis*[1,2,3] tmarghet@indiana.edu
Minje Kim[4] minje@indiana.edu

* lead authors contributed equally [1] Cognitive Science [2] Psychological & Brain Sciences
[3] School of Public & Environmental Affairs [4] Intelligent Systems Engineering
Indiana University, Bloomington, Indiana, USA

Abstract

High-level cognition is often accomplished not by individuals working in isolation, but by distributed, complex cognitive systems. Examples include teams of scientists or collaboratively improvising musicians. These distributed systems can undergo critical transitions, suddenly moving from one stable pattern of activity to another. For instance, in 'free jazz,' where musicians improvise without a predetermined plan or a central leader, the performance will often settle into a particular texture or style before transitioning to something entirely new, often quite suddenly. When do these transitions occur? Are they foreseeable? Inspired by suggestions that cognitive systems are, in some sense, a kind of 'ecosystem,' we draw on recent work in quantitative ecology that has begun to describe generic early warning signals of impending critical transitions in ecosystems. We apply these techniques to a corpus of audio recordings of professional jazz quartets playing improvised music. We find that the same generic measures that have been used successfully to predict critical transitions in natural ecosystems describe the complex dynamics of improvised musical performance in the lead-up to transitions. By taking seriously the metaphor that cognition occurs in 'ecosystems,' we gain new insights into how stable patterns of thought can emerge suddenly in complex cognitive systems.

Keywords: Early Warning Signals; Music; Improvisation; Complex Adaptive Systems; Distributed Cognition

Introduction

A massive team of scientists, entangled with highly advanced instruments, make a breakthrough discovery in particle physics. Sailors on a navy vessel creatively reconfigure the way they track the ship's location (Hutchins, 1995). Improvising musicians explore one particular musical texture before suddenly transitioning to a new style or sound. In each of these examples, expert reasoning is accomplished not by individuals working in isolation, but by the whole distributed, complex system. Such distributed cognitive systems can undergo critical transitions — sudden transformations from one stable pattern of activity to another. But when? Are these transitions foreseeable?

Consider two examples. First, Hutchins (1995) describes how sailors on a large navy ship were forced to improvise when the ship suddenly experienced a catastrophic loss of power. The typical procedure for tracking the ship's location relied on instruments that were no longer operational. The entire distributed system of sailors and instruments first stabilized around one procedure for locating the ship on a map; after it became apparent that this procedure was incorrect, the entire system reconfigured quite suddenly into a new distributed procedure.

As a second example, consider 'free jazz' musicians who create new improvised music without a score, a central leader, or explicit advance planning. Free jazz often exhibits unexpected transitions between qualitatively different musical textures characterized by stable rhythmic, melodic, or sonic frameworks — which we will call 'soundworlds.' After exploring a soundworld for some time, a jazz ensemble may transition to an entirely new soundworld, often quite suddenly, without apparent warning. These critical transitions between soundworlds appear to reflect the emergence of a new stable regime within the distributed system of musicians.

Both these examples illustrate how distributed cognitive systems can suddenly transition from one stable regime to another — resulting in entirely novel cognitive procedures or products. While the existence of these critical transitions is well known, we know little about why and when these transitions occur.

An ecosystems perspective

Distributed cognitive systems often involve a heterogeneous mix of agents that interact across space and time. Inspired by this, a number of authors have proposed that we adopt an 'ecological' perspective on distributed cognitive systems (Bateson, 1972; Gibson, 1986; Hutchins, 2010). This analogy between cognitive systems and ecosystems has largely remained at the level of an evocative image. Over the last decade, however, ecologists have begun to develop mathematical tools for analyzing, and perhaps even predicting, critical transitions in natural ecosystems.

While many natural ecosystems exhibit considerable resilience, maintaining a stable regime of activity in the face of outside perturbations, they can sometimes transition suddenly to an entirely different regime. A fish population may collapse suddenly, or a lush ecosystem may experience sudden desertification. These 'tipping points' can seem to appear from nowhere, but ecologists have recently begun to develop *early warning signals* (EWS) of impending critical transitions (Sheffer et al, 2009; van Belzen et al 2017; Dakos et al, 2012; Wang et al, 2012).

In particular, many critical transitions are preceded by a period in which the ecosystem becomes increasingly sensitive to perturbations. In response to some outside 'push,' a resilient ecosystem will eventually return to a stable state — but in systems that are susceptible to a critical transition, this process of restabilization will take longer and longer. This decrease in the rate-of-return is known as *critical slowing down* (Fig 1B). Ecologists have identified a number of indices that an ecosystem is undergoing critical slowing, including two that we will deploy here: increased variability, and increased lagged autocorrelation (Fig 1B).

From forest ecosystems to cognitive ecosystems

Can we gain traction on critical transitions in distributed *cognitive* systems by taking seriously the metaphor that distributed cognitive systems are a kind of 'ecosystem'? There are reasons to hesitate. The complex systems that accomplish human cognition often differ in critical ways from the ecological systems for which these generic early warning signals were first developed. Distributed cognitive systems often involve a large number of distinct roles — think of the proliferation of precise roles on a navy battleship, or dozen or more musical parts in a piece of modern orchestra music. Cognitive systems can also be highly heterogeneous, combining highly specialized humans with a variety of technologies, tools, and practices. And cognitive systems operate on multiple timescales, some of which on multiple timescales that are of less importance in natural ecosystems; in addition to evolution and moment-to-moment interaction, cognitive systems often have a *cultural history.* Jazz musicians, for instance, spend years acquiring a set of shared intuitions and practices.

To investigate whether early warning signals of critical transitions in natural ecosystems might also predict transitions in expert cognitive systems, we focused on a system that exemplifies many of the most rarified and unusual features of human cognitive activity: ensemble free jazz performance. Free jazz performance involves multiple humans in hyper-specialized roles, using a range of technologies (i.e., instruments), with skills honed over a lifetime, to create new music without an organizing score. The collective activity of this distributed system creates entirely novel musical products.

In the current study, using a corpus of audio recordings of professional jazz musicians, we quantify the amount of critical slowing down in the periods leading up to improvised transitions. If this class of distributed cognitive systems behaves similarly to natural ecosystems, then critical slowing down should increase systematically in the lead-up to a critical transition.

Methods

Corpus of free jazz recordings

The corpus consisted of single-track audio recordings of improvised free jazz music (~ 20 minutes), taken from two recording sessions in a professional recording studio. Both sessions featured a professional jazz quartet consisting of saxophone, guitar, bass and drums. Although the tracks were recorded on separate occasions, they feature the same musicians except for the saxophonist.

The first piece (length: 16m12s) was entirely 'free' improvised, making no use of pre-composed material and with no explicitly stated preconceptions about what was supposed to happen.

The second piece (length: 2m54s) included some composed material, interleaved with periods of free improvisation. Transitions between improvised and composed sections were determined collectively in the moment, so we consider these to be improvised. The timing of transitions between composed and improvised material were determined explicitly by the score, so we consider these to be composed.

Identifying critical transitions

From the audio recording, a professional musician coded the onset of transitions between soundworlds. Two critieria were used to identify critical musical transitions: (1) transitions needed to demarcate qualitatively different musical textures or styles, and (2) transitions needed to involve a coordinated change in playing from two or more individuals. The coder identified twelve transitions: 8 improvised and 4 composed. Only improvised transitions were included in our analysis.

Analysis

The audio recordings were transformed into a multi-dimensional time series using Mel-Frequency Cepstrum Coefficients (MFCC), which have been widely used in speech recognition and music classification (Rabiner et al, 1993; Tzanetakis et al, 2002). The typical procedure to calculate MFCCs is to (a) apply a sliding window to a short period of the signal, (b) transform the windowed signal to the frequency domain by using Discrete Fourier Transform, (c) take the log of the power of the spectrum, (d) convert the linearly spaced frequencies into the mel scale, (e) apply Discrete Cosine Transform, (f) move on to the next frame. MFCCs provide a spectral analysis whose small number of coefficients (usually 13) contains a compressed version of the full Fourier spectral analysis results. On top of that, this compact representation is more suitable than the original time domain signal with redundantly high temporal resolution. To further reduce the dimensionality of our data, each dimension of the MFCC time series was normalized before performing Principal Component Analysis. For simplicity, we used only the first four PCA components (Fig1, top).

From this four-dimensional time series representation of the musical performance, we computed two indices of critical slowing down: variability and lag-1 autocorrelation. Each was computed in a sliding window (17.5s), with a stepsize between windows of 0.02 seconds. Variability was computed as the distance between all points in a given window, and autocorrelation was computed as the mean lag-1 autocorrelation for each of the 4 components. This produced two time series of potential early warning signals; one for variability, and one for autocorrelation (Fig2, bottom).

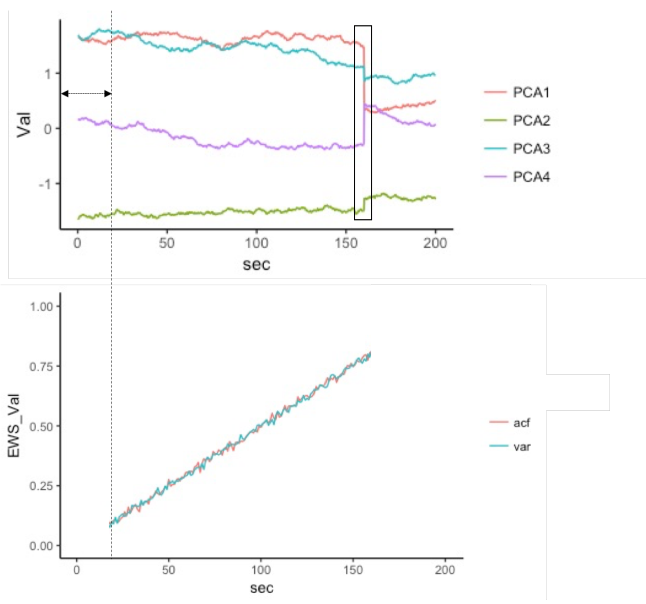

Figure 1. Audio recordings were transformed into a multidimensional feature representation, which was used to calculate putative early warning signals of critical transitions. (Top) Each single-track audio recording was transformed into a multidimensional feature representation (MFCC), then reduced to the first four principal components (colored lines). When performers transition from one soundworld to another (vertical rectangle), the performance should move to a new region of this 4-dimensional state space. (Bottom) From this 4-d representation of the audio, we calculated two putative early signals: variability and lagged autocorrelation. These were computed within a sliding window (length = 17.5 seconds), illustrated by the horizontal arrow in the top panel. In this toy example, both variability and lagged autocorrelation increase monotonically in the lead up to the critical transition.

Critical slowing down is indexed by a systematic increase in variability and lagged autocorrelation. We thus examined the rate of change in these early warning signals over the course of the entire performance, in a sliding window of length 35s and stepsize 0.02s. To quantify the rate of change within these windows we used Kendall's tau to measure the correlation between the early warning signal and time. More positive values of tau indicate greater monotonic increase. A system undergoing critical slowing down during a particular time period (e.g., leading up to a musical transition) should thus have a positive value of tau for this period.

Results

Soundworlds: sub-regions of musical state space?

We first verified that the multidimensional feature representation of the audio allowed us to distinguish between listener-perceived soundworlds. We thus investigated whether distinct soundworlds inhabited distinct (though perhaps overlapping) regions of the multidimensional feature space. A soundworld that consists of entirely the same sound played over and over should exist within a small region of the multidimensional feature space; conversely, moments from two very different soundworlds should be far apart in the feature space. Following this logic, for each soundworld, we found the subset of points that fully enclosed all points in the entire soundworld (i.e., the convex hull) and calculated the volume of this region. We then compared the size of each soundworld-region to the size of equally-sized random samples from the entire performance (n = 100 random samples per soundworld).

Each improvised soundworld inhabited smaller regions than expected by chance (every $p < .01$). Thus, our multi-dimensional feature space captured the subjective judgments of the listener, correctly placing moments from within a soundworld closer together than to other moments from the performance.

Critical slowing down before musical transitions?

We next turned to our critical question: Whether critical musical transitions were preceded by critical slowing down. As predicted, both indices of critical slowing down — variability and lagged autocorrelation — increased systematically and selectively in the period leading up to a critical transition. These warning signals were increasing monotonically, as measured by Kendall's tau, at most critical transitions between soundworlds, for both variability (75% of transitions; $M = 0.24 \pm 0.19$ SEM) and for lagged autocorrelation (75% of transitions; $M = 0.28 \pm 0.17$ SEM). By comparison, for instance, these early warning signals were largely flat during the rest of the second half of the soundworlds (variability: $M = 0.05 \pm 0.14$ SEM; autocorrelation: $M = 0.05 \pm 0.17$ SEM). The complex systems generating these soundworlds, therefore, appeared to undergo critical slowing down in the lead-up to transitioning between soundworlds.

To account for non-stationarity, we compared these results to surrogate time series with the same linear trend but scrambled noise structure. We fit a linear model to the original multidimensional feature representation of the audio and randomly scrambled the residuals across time, creating a surrogate time series that controlled for drift in the data while randomizing the temporal spread of the noise structure. Comparing the early warning signals in the original time series to equivalent times in the surrogate time series allows us to rule out the possibility that increases in early warning signals are driven entirely by non-stationarity.

We analyzed the timecourse and specificity of critical slowing down with a linear mixed effects model of the *tau* values across time. This model had fixed effects of soundworld-normalized time (start = 0, end = 1), the time series (actual = 0, surrogate = 1), and their interaction, along with random by-soundworld intercepts and effects of time series (actual vs. surrogate).

This model confirmed that both variability and lagged autocorrelation increased monotonically in the lead-up to a critical transition (variability: $b = 0.43 \pm 0.08$ SEM, $p < 0.001$; autocorrelation: $b = 0.30 \pm 0.09$ SEM, $p = 0.02$). This increase was damped significantly in the surrogate time series (variability: $b = 0.32 \pm 0.14$ SEM, $p = 0.047$; autocorrelation: $b = -0.32 \pm 0.12$ SEM, $p = 0.04$). Finally, critical slowing down increased significantly over the course of the soundworlds (variability: $b = 0.77 \pm 0.15$ SEM, $p < 0.001$; autocorrelation: $b = -0.47 \pm 0.18$ SEM, $p < 0.001$) — but this occurred primarily in the actual time series rather than the surrogate time series (variability: $b = -0.56 \pm 0.02$ SEM, $p < 0.001$; autocorrelation: $b = -0.40 \pm 0.02$ SEM, $p < 0.001$). In summary, these distributed systems underwent critical slowing down in the period leading up to a critical transition.

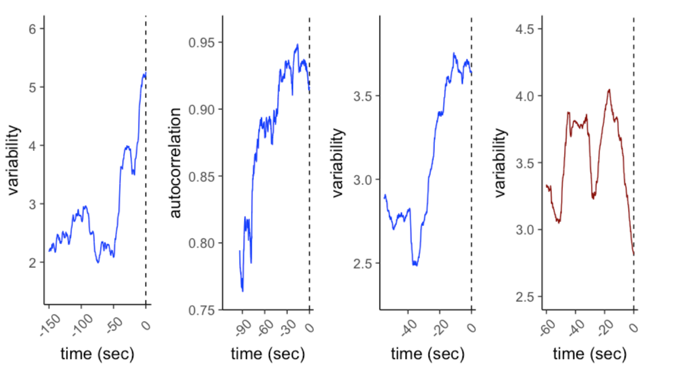

Figure 2. Early warning signals increased systematically before critical transitions. The first three plots (blue) show soundworlds where the early warning signal (vertical axis) increased systematically leading up to a musical transition (dashed lines). For comparison, the fourth plot (red) shows a soundworld where the performance did not exhibit critical slowing down before transitioning (i.e., variability decreased, rather than increased). The predicted increase in the early warning signals before a transition is captured by the correlation between the early warning signal and time (i.e., Kendall's tau), plotted in Figure 3.

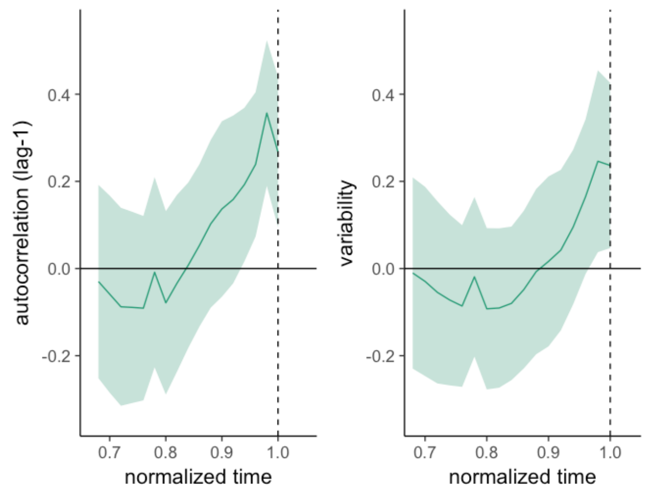

Figure 3. Mean critical slowing down in the period immediately before a musical transition (dashed line). The vertical axis indicates the direction and rate of change (i.e., Kendall's tau) for each early warning signal in a 35-sec. sliding window ending at that moment. Positive values indicate monotonic increase. Time (x-axis) is normalized to go from 0 to 1 within each soundworld. Error ribbon = standard error of the mean.

Discussion

We asked whether adopting an 'ecosystem' perspective might give us empirical traction on an outstanding problem: understanding when and why critical transitions occur in distributed cognitive activity. As our case study, we chose the sudden musical transitions that occur in improvised free jazz performance. Building on theoretical and empirical investigations of critical transitions in natural ecosystems, we calculated generic early warning signals that are thought to index critical slowing down, a kind of loss-of-resilience that can precede critical transitions. As predicted, we found that improvised transitions between musical 'soundworlds' were preceded by a monotonic increase in two measures of critical slowing down: variability and lagged autocorrelation. In free jazz ensembles, sudden transitions from one distributed cognitive regime to another appear to follow the same general patterns that characterize transitions in natural ecosystems.

Improvised music as a cognitive ecosystem

Some of the most sophisticated and intricate formal systems invented by humans have arisen in the domain of music: tonal harmony of Western classical music, the polyrhythms of African drumming, melodic counterpoint of the Bach inventions. In composed music, musical structures (i.e. the surface realizations of these formal systems) are formulated ahead of time by an individual composer and dictated to performers in the form of a written score. But in improvised music, abstract musical structures emerge spontaneously out of the distributed activity of the ensemble. Interaction is key in improvised music. Each individual is continuously adapting to and influencing the other members of the ensemble (Walton et al, 2015). Musical meaning is less ascribable to individual intentions as it is to the ongoing interplay between the various voices of the ensemble (Borgo, 2005). Improvised music thus provides us a testbed to study how high-level cognition (in the form of abstract musical structures) emerges out of distributed action of embodied, highly specialized agents.

There are many genres of improvised music across the globe. Each musical tradition can be characterized by distinct conceptual structures, functional norms and roles aesthetic guidelines. For example, the raga in Indian classical music provides a framework for improvising coherent melodies over long song forms. Jazz musicians master a shared repertoire of "standard" tunes – melodies with corresponding harmonic structures, that serve as improvisational templates. Musical genres also determine functional roles assigned to particular instruments. In straight ahead jazz for instance, the bassist typically "walks" (i.e. plays quarter notes to mark time and passing harmonic structure) while the saxophonist improvises a melodic solo over the underlying harmony. These culturally construed functional roles and formal structures are learned in intimate detail by improvisers ahead of time, whose shared knowledge facilitates the spontaneous generation of sophisticated, compelling musical pieces. The existence of these culturally produced constraints, and of learning on the part of individual musicians constitutes an interesting departure of cognitive ecosystems from natural ecosystems.

In this paper we analyzed recordings of free jazz. In free jazz, musicians with strong training in straight-ahead jazz come together to improvise without reference to any preconceived "tune" or template. The musicians have a shared mastery of the formal structures of straight-ahead jazz, but they are not confined by them. Functional roles can be challenged, new systems of harmony and rhythm can be explored. In the course of this exploration, groups often settle into stable regimes, which can be characterized by distinct harmonic structures, rhythmic patterns, or sets functional roles. For example, in the quartets analyzed in this paper, there might be a cacophonous section in which everyone is playing as loud and fast as possible with little to no group coherence. This section may then yield to an intimate duo with saxophone and guitar carefully co-constructing melodic and harmonic material.

Most of the periods leading up to transitions between different stable regimes showed evidence of critical slowing down. This is due to the collective, decentralized manner in which those transitions are executed. Within a given stable regime, one performer within the ensemble may hint at a new musical area to explore. This hint might be reinforced by another musician playing a supporting motif, which might be further reinforced by a third musician catching onto the developing theme. In this manner, an ensemble can quickly transition between qualitatively different stable regimes without any central locus of control. While this may be the norm, it is important to note that improvising ensembles also have the capacity to make centralized transitions. In some cases, an individual may simply decide to start playing something different and force the ensemble along an alternate path. An improvising ensemble's capability to support both decentralized and centralized dynamics constitutes an interesting distinction with natural ecosystems.

Future work

In future iterations, we would like to use a more sophisticated modeling technique in our surrogate analysis. Here drift was modeled with a linear regression spanning the entire range of the recording, but it is possible that shorter-spanned nonlinear trends present in our dataset were not captured by this approach. Following the example of past works, it may be beneficial to fit an ARIMA model to the music timeseries (Wang et al, 2012).

Another extension of this work will be to add more recordings to the corpus. Doing so will enable increased statistical confidence in the results, as well as an opportunity to analyze the behavior of different musical ensembles. Moreover, adding composed music to our dataset would enable comparison between improvised versus composed transitions. If it is true that the CSD observed in the free jazz transitions owes to the distributed nature of improvised performance, we should expect to *not* observe CSD in composed transitions as they are issued by an *a priori* script (i.e. the written score).

Conclusion

Much high-level cognition is accomplished by systems that are complex, distributed, and adaptive. Political consensus requires multiple individuals bringing their beliefs into alignment. Scientific activity is almost always a community endeavor. Great musical improvisation often requires each performer to cede some autonomy to the emergent will of the group. Each of these systems can stay in a particular regime for a prolonged period of time — before suddenly transitioning to a new cognitive state. One political consensus might break down, replaced by another. Scientists have *eureka* moments. Musicians shock their audiences — and themselves — by playing something that has never been played before. Describing *when* these critical transitions occur is a first step towards understanding why.

Acknowledgments

We thank members of the GeoLab for comments.

References

Bateson, G. (1972). *Steps to an ecology of mind*. New York: Balentine Books.

Borgo, D. (2005). *Sync or Swarm: Improvising Music in a Complex Age.* Bloomsbury Academic.

Belzen, J. V., et al. (2017). Vegetation Recovery in Tidal Marshes Reveals Critical Slowing down under Increased Inundation. *Nature Communications*, *vol. 8.*

Dakos, V., et al. (2012). Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data. *PLoS ONE*, *vol. 7, no. 7.*

Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum.

Hutchins, E. (1995). *Cognition in the Wild.* MIT Press.

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, *vol. 2, no. 4*, 705–715.

Rabiner, L. R., & Juang, B.-H. (1993). *Fundamentals of speech recognition.* Englewood Cliffs: PTR Prentice Hall.

Sheffer et al. (2009). Early-warning signals for critical transitions. *Nature*, *vol. 461.*

Tzanetakis, G., and Cook, P. (2002). "Musical genre classification of audio signals." *IEEE Transactions on speech and audio processing*. 293-302

Walton, A., et al. (2015. Improvisation and the Self-Organization of Multiple Musical Bodies. *Frontiers in Psychology*, *vol. 06.*