This is the Remix: Structural Improvisation using Automated Pattern Discovery

Sean Whalen
Department of Computer Science
Columbia University
New York, NY 10027
shwhalen@gmail.com

James P. Crutchfield
Complexity Sciences Center
Physics Department
University of California, Davis
Davis, CA 95616
chaos@ucdavis.edu

1 Introduction

The remix is pervasive in modern culture. Popularized in contemporary art by early hip-hop, it traces its roots back to both Dadaism and, later, the cut-ups of William S. Burroughs [1]. The basic approach is well understood: Take pre-existing material, then subtract, add, and adjust until something novel and engaging emerges. The remixer becomes a mash-up unto themselves: part sculptor, part curator, part engineer.

In popular music, the remix attempts to inject freshness into a familiar song—perhaps to make a slow song danceable or to blend the vocals of a song with the rhythms of a disparate genre in a way that creates new fans. In contrast, the cut-ups of Burroughs or the randomized piano fragments of Stockhausen's *Klavierstücke XI* deny such familiarity. And, as a result, they require more of the reader or listener. In all of these, though, the interplay between structure and randomness is the very foundation of our cognitive appreciation: a perceptual information channel from creator to consumer in which both total order and total chaos become uninteresting. Between these two extremes exists an ideal dynamic trade-off where familiarity and surprise keep our attention [2]—when we experience the emergence of structural complexity.

2 Structural Complexity: Between Order and Chaos

We now understand the mathematical underpinnings of how interestingness arises in the middle ground between order and chaos [3]. In short, there are two basic characters in a signal source. The first is how random it is. As observers, we experience a certain degree of unpredictability or surprise. We make errors in anticipation of its future. The second character is how structured or patterned the source is. How much memory is used to produce a given degree of randomness? In what architecture is that memory stored?

Recent theoretical results demonstrate that these two characters are not tautological complements of each other—that "patternedness" is not the opposite of randomness [4]. Randomness is most directly measured using Shannon's entropy rate which captures the degree of surprise [5]. While the amount of pattern is measured by the source's statistical complexity which measures the amount of the past that is relevant to the present and to minimizing surprise. In short, entropy and complexity are independent coordinates for the space of signals [6].

There are many, many ways to estimate a signal's entropy rate; with the rate of compression via the Lempel-Ziv algorithm being probably the most accessible (via `gzip`) [7]. The statistical complexity, by contrast, is uniquely estimated by reconstructing the signal's minimal causal representation—the signal's ϵ-machine [8]. (Fortunately, the entropy rate is also efficiently estimated from the latter.)

The net result is that one discovers the patterns (causal states) in a signal using the signal itself. We call this procedure *pattern discovery* to distinguish it from pattern *recognition* algorithms that assume a priori what the relevant patterns are.

3 Auto-Remix

We have adapted these new insights to the domain of music synthesis. We give the job of remixer to a new entity—the song itself. Drawing on techniques for automated pattern discovery, we infer a signal's ϵ-machine—its minimally sized, maximally predictive stochastic model [8]. The resulting *Auto-Remix* algorithm is as follows:

- Divide the song into fragments. Here, we use beats as fragments to maintain the flow of the original song, though granules or other structures can be used.
- Cluster the fragments into groups based on their sonic similarity using the k-means algorithm.
- Create a time series of cluster identities based on the cluster membership of each fragment.
- Infer a generative state machine model (ϵ-machine) from this time series, grouping predictively equivalent histories into causal states.
- Generate a new song from a causal-state-to-causal-state random walk through the paths allowed in the ϵ-machine, emitting an audio fragment at random from within the cluster associated with each causal state along the path.

Auto-Remix allows the song itself to become its own source material at the level of beats, granules, or other levels of abstraction. A preference for familiarity or surprise is expressed during the clustering process, between the extremes of assigning each fragment its own cluster or grouping many (possibly dissimilar) fragments into the same cluster. The time-ordered structure of these fragments is then inferred by an automated discovery algorithm, producing the ϵ-machine—a special class of Hidden Markov Model (HMM) whose topology and transition structure are inferred directly from the statistics of the data. The resulting ϵ-machine is a variable-order HMM that allows for a natural trade-off between order and chaos through both the clustering process and the amount of historical information used to infer its state-based representation. The original song expresses itself at various points along this complexity-entropy spectrum, from Cage and Xenakis (disordered) to Stockhausen (semi-ordered) and back to its original form (ordered).

4 Perceptual Metrics

Due to distinct features (minimality and unifilarity) of the ϵ-machine representation of the signal, during Auto-Remix one has immediate access to key signal properties and so a baseline for their perception. The amount of temporal correlation, how much of the past is currently relevant, is measured via the instantaneous statistical complexity. The amount of surprise, the degree of information generation, is estimated as the instantaneous Shannon entropy rate. These directly reflect where along the order-disorder spectrum one is and how structured the song can appear. Monitoring signal sources in such a complexity-entropy space has proven valuable in understanding the very organization of the space of signals [6]. And, based on our experience to date, we conjecture that in time Auto-Remix will come to provide access to these psychophysical metrics in the domain of music. As such, it will be grounded in aural-psychophysics and a useful companion to aesthetic and syntactic approaches, such as found in [9].

5 Concluding Remarks

The Auto-Remix algorithm infers a stochastic model of a song at different levels of abstraction and allows the song's intrinsic structure to re-interpret itself. Using automated pattern discovery techniques in combination with clustering, a new interpretation can be expressed anywhere in the spectrum of complexity from complete order to complete chaos. The cognitive relevance of such a range was recently demonstrated. The brain's reward cues are stimulated by signals that tend toward the structured, but maximized by incorporating elements of surprise [2, 10]. Auto-Remix not only allows one to explore the aesthetics of these tradeoffs, but also to quantify them using information-theoretic metrics derived from our stochastic model—the ϵ-machine.

References

[1] Robin Lydenberg. *Word cultures: Radical theory and practice in William S. Burroughs' fiction*. University of Illinois Press, 1987.

[2] Valorie N. Salimpoor, Mitchel Benovoy, Kevin Larcher, Alain Dagher, and Robert J. Zatorre. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, 14(2):257–262, 2011.

[3] James P. Crutchfield. What lies between order and chaos? In J. Casti and A. Karlqvist, editors, *Art and Complexity*, pages 31–45. Elsevier Science B.V., 2003.

[4] J. P. Crutchfield. Is anything ever new? Considering emergence. In G. Cowan, D. Pines, and D. Melzner, editors, *Complexity: Metaphors, Models, and Reality*, volume XIX of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 479–497, Reading, MA, 1994. Addison-Wesley. Santa Fe Institute Technical Report 94-03-011.

[5] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, second edition, 2006.

[6] D. P. Feldman, C. S. McTague, and J. P. Crutchfield. The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing. *CHAOS*, 18(4):043106, 2008.

[7] J. Ziv and A. Lempel. Compression of individual sequences via variable-rate encoding. *IEEE Trans. Info. Th.*, IT-24:530, 1978.

[8] J. P. Crutchfield. The calculi of emergence: Computation, dynamics, and induction. *Physica D*, 75:11–54, 1994.

[9] D. Cope. *Virtual Music: Computer Synthesis of Musical Style*. The MIT Press, Cambridge, Massachusetts, 2001.

[10] Leonard B. Meyer. *Emotion and meaning in music*. University of Chicago Press, 1961.