Analysing and modelling complex rhythm dynamics in self-organising entities using Stigmergy with application to more lifelike automated intricate rhythm pattern generation

Joseph Timoney
jtimoney@cs.nuim.ie

Puya Amiri
puya.amiri@mail.polimi.it

June 2018

As rhythms in a single song interchange between isochronous, metrical, and non-metrical rhythm within seconds, the brain of an experienced drummer has been found to respond instantaneously. Isochrony is an essential characteristic of human drumming. It means evenness that is, making beats occur at regularly spaced time intervals so that the occurrence of the next beat is predictable. The evenness facilitates temporal coordination among percussionists and is a remarkable feature of human music. Additionally, every other percussive gesture inside of this interval is felt to be a subset of this greater pattern, embellishing it, adding a sensation of movement. It is the human feel though that makes the sequences of the other percussive notes possess what is sometimes referred to as 'groove', a magical quality. Also a group of players are called "Being in the groove" when they show an advanced level of synchronous pattern development. It is a cognitive temporal phenomenon emerging from one or more carefully aligned concurrent rhythmic patterns, characterized by perception of recurring pulses, and subdivision of structure in such pulses,perception of a cycle of time, enabling identification of cycle locations, and effectiveness of engaging synchronizing body responses (e.g. dance, foot-tapping). In contrast, a percussion line that is played exactly as written sounds rigid and mechanical. Our exposure to electronic dance music has illustrates this contrast clearly, there is no mistaking the 'feel' of a real drummer from the machine. Is it possible to capture the artistic expression and cognitive processes of a human performer or performers and obtain a more vivid interpretation of a percussion performance by a computer? To achieve this rhythms must be played with an appropriate feel expressed though subtleties in the micro timing and the hit strength or velocity. Furthermore, according to Temperley (Temperley, 2001)[10] listeners match structures in the percussive line against sound events in a synchronization and lay a foundation for the further perceptual organization

There have been many attempts to generate algorithmic music that has a natural feel using techniques that includes Markov models, generative grammars, genetic algorithms, and neural networks. The problem becomes much harder too when polyphony, that is multiple lines of music that occur simultaneously, is involved. The exact same is true for percussion, as with a drum kit or percussion section the 'groove' is derived from how the patterns across the instruments interact. It is only very recently that researchers are attempting to tackle the difficulty of polyphony with musical lines. The approach is to extend what works for single lines. For example, Recurrent neural networks (RNN), especially long short-term memory networks (LSTM) have been shown to be extremely effective at modelling single-dimensional temporal patterns. Using them to model polyphonic music, a naive approach is to treat all of the notes played at any given timestep as a single input vector, and train an LSTM network to output a vector of probabilities of playing each note in the next timestep. This essentially models each note as an independent event. However, real-world polyphonic music contains complex relationships that are better described using joint probability distributions. To this end, a more effective approach creates a structure that combines RNNs and restricted Boltzmann machines to model the joint probability distribution of notes at each timestep. One work specifically examines two versions: Tied Parallel LSTM-NADE (TP-LSTM-NADE) and Biaxial LSTM (BALSTM). Both models were found to give more emphasis to recent notes over the long-term structure. For drumming specifically another study examined two types of recurrent neural networks (RNNs) for modelling: stacked Long-Short Term Memory (LSTM) RNNs, and Clockwork (CW) RNNs. It was found that as a generative model the neural

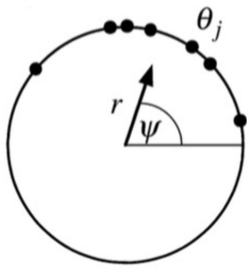

Figure 1: Kuramoto, Complex order parameter

networks were able to pick up on structural features of the beat onsets in the training set and generate them with some amount of variability in the output. Additionally, the network might be able to learn patterns of activation between different percussion instruments in the set as opposed to relying on a one-dimensional representation of pulse.

This leads to another important aspect of interaction modelling. In multi-object rhythmic analysis the concept of Entrainment is prevalent. Entrainment refers to a one-sided interaction where one rhythmic object, modelled as an oscillation, acts as a slave and synchronises with another oscillator, termed the master. A well-known mathematical mode for this type of behaviour, which strikes a compelling balance of generality and solvability, was proposed by Kuramoto in 1975. In such a system the behaviour and bulk properties of a set of slave oscillators synchronizing to a master is described in terms of their phases and a coupling parameter. Each oscillator is described by a time-dependent phase [1] $\theta(t)_i$, which in the absence of coupling rotates at its natural frequency $\omega(i)$.

$$\dot{\theta(t)}_i = \frac{\mathrm{d}\theta_i}{\mathrm{d}t} \tag{1}$$

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) \tag{2}$$

N = Number of oscillators
$\omega(i)$ = initial frequency
K = coupling coefficient
θ_j = phase of jth oscillator
θ_i = phase of ith oscillator

Equation2 can be written in terms of a complex order parameter, $r(t)$. This can be thought of as the "Collective rhythm" that is produced by whole oscillator population. r is the vector radius of the circle that moves at a frequency of ψ. $[r, \Psi]$ (t) enable a macro-level view of the model:

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta j} \tag{3}$$

$r(t)$ is a good indicator of phase coherence: how well the oscillators are locking into the same phase. $\Psi(t)$ indicates average group velocity, figure1.

However, in a more complicated synchronised oscillation system, there is a degree of mutual feedback where the oscillators find a common frequency and phase. This synchronisation is a self-organising phenomenon unlike with Entrainment, as strictly speaking in this case the master oscillator centrally controls the system of oscillators. One new approach to synchronization has been termed Stigmergy. This models how oscillating entities are indirectly influenced by each other by adopting the notion of local field coupling and includes the external effects of noise, distance, delay and influence. This new model is different to the Kuramoto model. One method for describing stigmergy is exploiting Van der Pol oscillators, equation 4 shows system of two van der Pol

[1] in some literatures it is named instantaneous phase

oscillators x and y, coupled via a "bath" z.

$$\begin{aligned} \ddot{x} - \epsilon(1 - x^2)\dot{x}^2 + x &= k(z - x) \\ \ddot{y} - \epsilon(1 - y^2)\dot{y}^2 + y &= k(z - y) \\ \dot{z} &= k(x - z) + k(y - z) \end{aligned} \tag{4}$$

However, the mathematical development for the Stigmergic system is theoretically undeveloped. Therefore, the idea for this research is to create a new rhythmic generator technology. This system will be capable of generating multiple lines of percussion that sound natural. The synchronization will be self-organising and thus will be governed by a Stigmergic model that must be derived. The rhythm generator based on this Stigmergic system will be suitable neural network technology that can manage both short-term and long-term patterns in the percussive line. Micro-timing modulation and velocity variations will be integral to the model to impart the sensation of 'groove'. An outline for the work is:

- Investigate and derive a suitable Stigmergic self-organising oscillator system model
- Investigate, Implement and test a variety of the neural network technologies
- Carry out extensive user testing
- Review results and feedback to model improvement
- Finalise models followed by validation testing

Aside from using this new technology for more natural musical drumming performances is there other application that could benefit from more human-like models of percussive interactions? In scientific studies it has been found that musicians have significantly faster reaction times and elevated rhythm-change detection performance versus non-musicians. Such observations have influenced the development of Music-supported therapy (MST) as an innovative and positive approach towards aiding brain function, particularly in the aged. Long-term training of cognitive function utilizing MST therapy can help develop an improved sense of pattern understanding that ultimately enhances focus, establishes regularity, and works toward restoring degrees of normal brain function. Most importantly, these positive changes improve quality of life: the main goal of MST.

References

[1] Joseph Colonel, Christopher Curro, and Sam Keene. Improving neural net auto encoders for music synthesis. In *Audio Engineering Society Convention 143*, Oct 2017.

[2] Ciesielski V. DCorreia, J. and A. Liapis. Computational intelligence in music, sound, art and design. *Cham: Springer International Publishing.*, 2017.

[3] Logan James Deyo. Cognitive functioning of drumming and rhythm therapy for neurological disorders. university of tennessee honors thesis projects. *University of Tennessee Honors Thesis Projects.*, 2016.

[4] Nicolas & Curé Charlotte Elisabeth & Sterck Elisabeth. Dufour, Valérie & Poulin. Chimpanzee drumming: A spontaneous performance with characteristics of human musical drumming. *scientific Reports. 5. 10.1038/srep11320*, 2015.

[5] A. Lambert. Stigmergic model for oscillator synchronisation and its application in music systems. *Proceedings International Computer Music Conference, Lubjana.*, 2012.

[6] Nolan Lem. A Generative Model of Pulse Percept: Analyzing Performances of Free Jazz Drumming using Dynamic Beat Tracking and Recurrent Neural Networks. 2018.

[7] Rainer Polak, Justin London, and Nori Jacoby. Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: A corpus study of malian jembe drumming. *Frontiers in Neuroscience*, 10:285, 2016.

[8] Andrea Ravignani and Guy Madison. The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, 8, Jun 2017.

[9] Stephen Strogatz. Sync: How order emerges from chaos in the universe, nature, and daily life, hyperion. 2012.

[10] D. Temperley. The cognition of basic musical structures. *The MIT Press. Cambridge, Mass.*, 2001.

[11] Oliver Weede. Polymetric rhythmic feel for a cognitive drum computer. *CoRR*, abs/1606.06197, 2016.