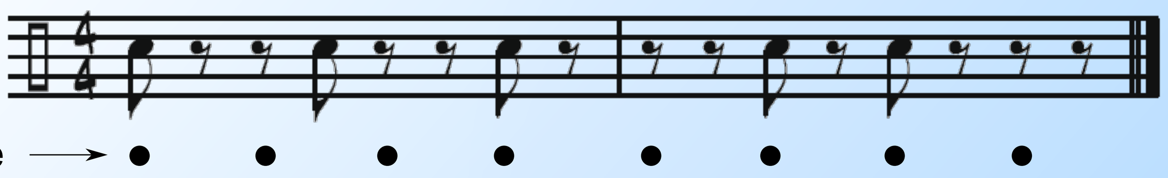
Deep Rhythms

Towards Structured Meter Perception, Learning and Generation with Deep Recurrent Oscillator Networks

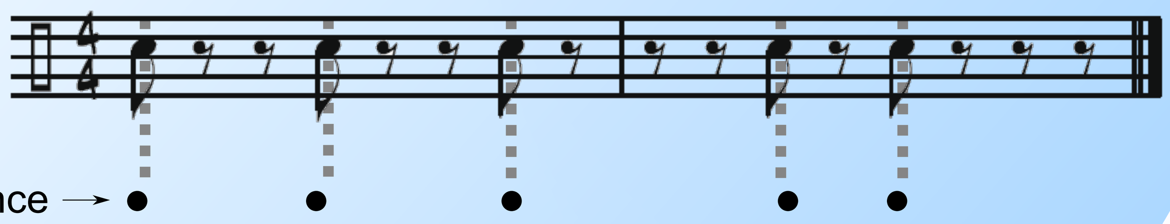
Andrew Lambert[12], Tillman Weyde[1], Newton Armstrong[2]

[1]Department of Computer Science, [2]Department of Music
City University London

Beat induction allows us to tap along to the beat of music, perceiving its pulse. This perceived pulse can be present in the stimulus, but it is often only implied by the musical events.

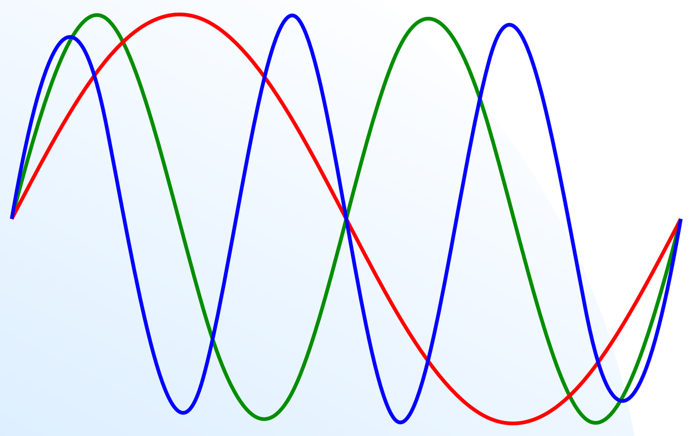


What's more, performed music is rarely periodic and subject to the performer's expressive timing.



This makes beat induction difficult to model computationally.

The process by which we can induce a beat is known as entrainment. Entrainment is the co-ordination of temporally structured events through interaction where two or more periodic signals are coupled in a stable relationship.

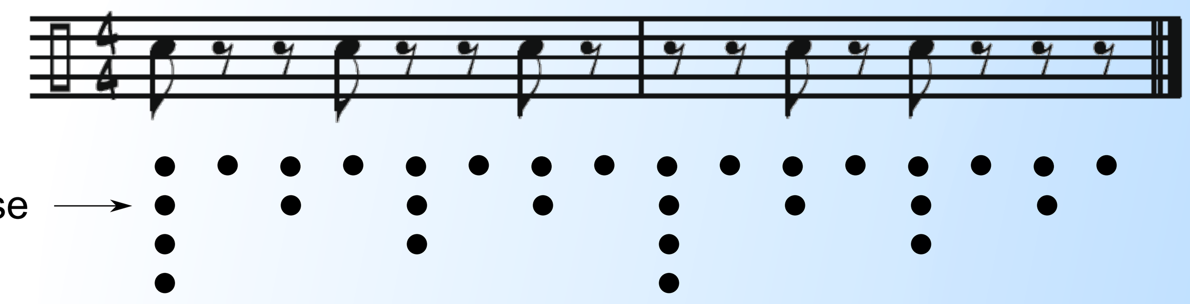


Snowball the dancing Cockatoo (Patel et al., 2009)

Ethnomusicologists are increasingly becoming aware of its importance as an approach to understanding music making and music perception as an interactive process (Clayton et al., 2005).

Finding the pulse of music is a step towards achieving other musical perception tasks, including meter perception.

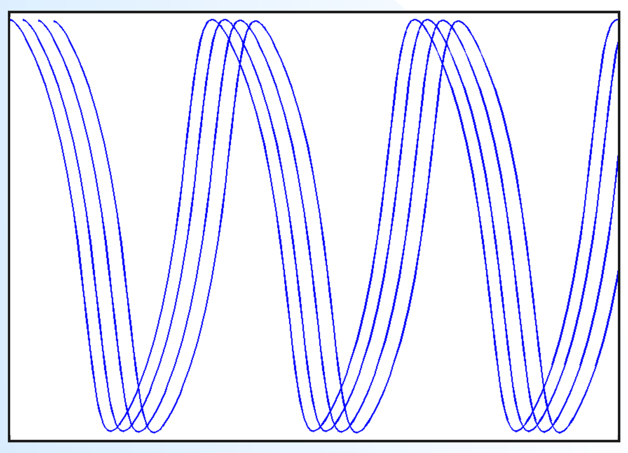
Meter refers to the multi-layered divisions of time present in music, of which the referent layer is the pulse.

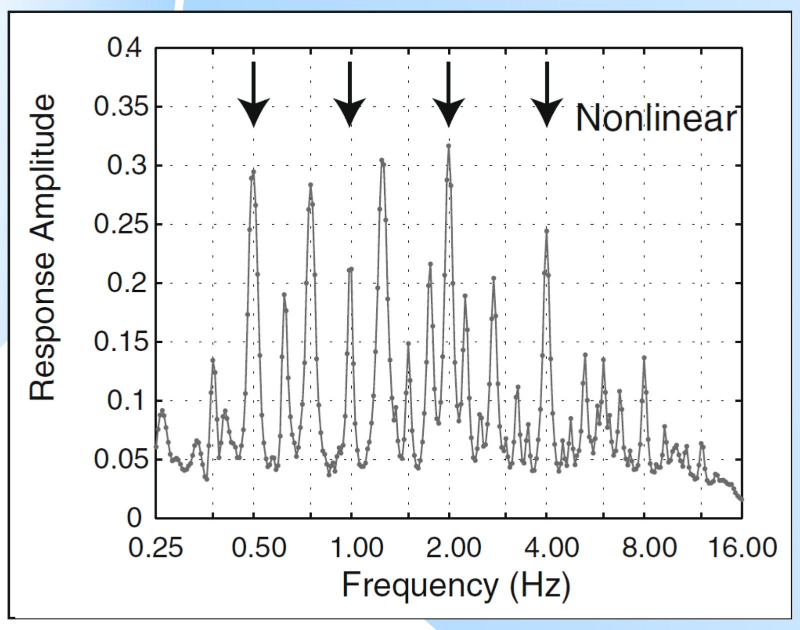


Other layers in music divide the pulse further to the smallest subdivisions of a beat, and extend towards larger measures, phrases, periods, and even higher order forms.

The individual components of music, the events in time, lead to the formation of new macroscopic spatial, temporal and functional structures in meter. These structures vary and repeat with time in their own patterns.

Nonlinear oscillator networks, modelled on neural oscillations, have been shown to achieve beat induction via their entrainment properties (Eck, 2002).





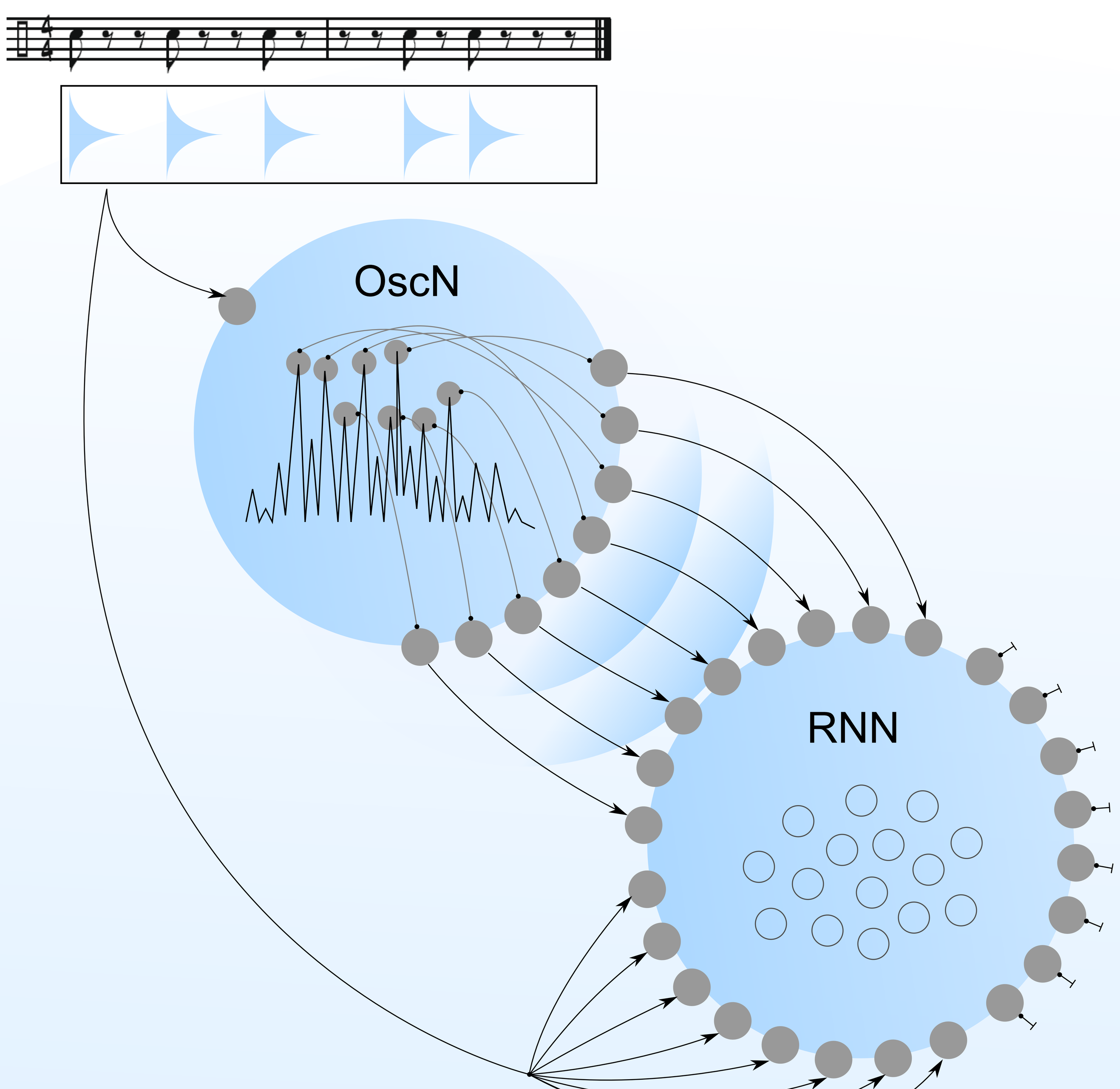
Furthermore, the phenomenon of nonlinear resonance, which arises out of a network of coupled nonlinear oscillators, actually adds frequency information to the signal. This can account for pattern completion, the perception of the missing fundamental, tonal relationships and the perception of meter (Large, 2010).

Most prior work on pulse and meter perception has been concerned with abstract temporal information, such as crafted pulses in time (Temperley, 2004). However meter perception develops through cultural learning and also relies on pitch.

This project aims to model these aspects of meter perception by designing a deep network.

At the core of the system will be a nonlinear oscillator network (OscN), which acts as an entrained resonant filter to the musical signal. The second part of the system will be a Recurrent Neural Network (RNN), which are known to have the ability to learn the kind of temporal structures required to perceive meter (Eck and Schmidhuber, 2002).

The system will operate on symbolic music in MIDI format.



In our design, the OscN has one input for the combined velocity of the signal, and eight outputs containing the strongest stable resonances. The RNN has eight inputs from the OscN in addition to one input and one output per pitch.

The OscN is driven in continuous time steps, whereas the RNN is driven in discrete time, quantized at the 16th beat level. This will allow the RNN to learn time scales relative to the perceived meter.

Learning is done by the gradient descent method Back Propagation Through Time. We will be using a training corpus of monophonic music, divided into genre.

After learning is complete the system will be seeded with a series of pitches and the RNN's outputs are connected back to the inputs, resulting in a generative model.

This work is the first experiment in a PhD project which aims to use networks of this type for music analysis and music information retrieval tasks, as well as music generation.

Clayton, M., Sager, R., & Will, U. (2005). In time with the music: the concept of entrainment and its significance for ethnomusicology. European Meetings in Ethnomusicology, 11, 3–142.
Eck, D. (2002). Finding downbeats with a relaxation oscillator. Psychological Research, 66(1), 18–25.
Eck, D., & Schmidhuber, J. (2002). Finding temporal structure in music: blues improvisation with LSTM recurrent networks. In Proceedings of the 2002 12th IEEE Workshop on Neural Networks for Signal Processing, 2002 (pp. 747–756).
Large, E. W. (2010). Neurodynamics of Music. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), Music Perception (pp. 201–231). Springer New York.
Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. Current Biology, 19(10), 827-830.
Temperley, D. (2004). An Evaluation System for Metrical Models. Computer Music Journal, 28(3), 28–44.