

A Network of Relaxation Oscillators that Finds Downbeats in Rhythms

Douglas Eck

Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA)
Galleria 2, CH 6928 Manno, Switzerland
doug@idsia.ch, <http://www.idsia.ch/~doug>

Abstract. A network of relaxation oscillators is used to find downbeats in rhythmical patterns. In this study, a novel model is described in detail. Its behavior is tested by exposing it to patterns having various levels of rhythmic complexity. We analyze the performance of the model and relate its success to previous work dealing with fast synchrony in coupled oscillators.

1 Downbeat Induction

The term *beats* refers to sounds that are perceived as being equally spaced in time. *Downbeats* are particularly salient beats that usually occur at a comfortable tapping rate. When you tap your feet to the radio you are finding downbeats, a skill called *beat induction*. Downbeats act as a unifying force, lending music the feeling of movement by allowing the listener to predict the onset of important musical events. The process of beat induction is influenced by many aspects of music including harmony, melody and rhythm [2, 9]. Because interactions are not always simple, it can be difficult to predict the locations of downbeats. For example, in rock and roll music, even though chord changes (harmonic components) usually occur on the first note of a musical bar, downbeats are usually aligned with the second note due to syncopated drumming style. Furthermore, although beats are *perceived* as being equally-spaced in time, perfect timing is rarely if ever found on the radio. Both motor noise and deviations due to expressive timing are found in performed music.

The task of beat induction can be simplified by considering only patterns of equal-amplitude beeps or clicks. This removes the influence of melody and harmony (and in addition amplitude and timbre variations). However, even in this “rhythm-only” domain, two important influences compete to decide downbeat location. First, the *grouping* of events in the pattern is important. For example, when three or more events are presented in rapid succession, the first and last events are more perceptually salient than the middle one. Second, the *meter* of the pattern is important. Meter is the sense of strong and weak beats that arises from interactions among hierarchical levels in a pattern having nested periodic components. Such a hierarchy is implied in Western music notation, where different levels are indicated by kinds of notes (whole notes, half notes, quarter notes, etc.) and where bars establish measures of an equal number of beats. For an overview see [5].

For these reasons, downbeat induction can be seen as a seemingly simple cognitive skill (tapping along with music) that is in fact challenging to model. We address this challenge by presenting a network of coupled oscillators that performs downbeat induction. The network consists of Fitzhugh-Nagumo oscillators [3, 13] designed to model the dynamics of neural action potential.

We show in this paper that they can also “tap along” when forced by rhythmical pulses of voltage. Computational simulations using a popular set of rhythmical patterns reveal that our model succeeds at finding the most likely downbeat in almost all cases. In the paper we analyze model performance and provide a possible explanation of why a small network of simulated neurons would do so well at the task.

In the literature, many have addressed the question of whether a simple model can find downbeats in patterns. Longuet-Higgins & Lee [10] and Povel & Essens [15] offer rules based symbolic models that find downbeats in symbolic patterns. Parncutt [14] considers the importance of presentation rate in beat induction. Large & Kolen [8] and McAuley [12] provide oscillator models that work online to find downbeats in temporal signals.

2 Model Details

The Fitzhugh-Nagumo oscillator is a two-variable caricature of the four-variable Hodgkin-Huxley model of neural action potential [6]. It exhibits the integrate-and-fire spiking dynamics of a real neuron. When fed a constant amount of low voltage, the oscillator gradually accrues its own voltage until it reaches a threshold; upon reaching that threshold it fires and quickly releases the energy. If the energy source is sufficient and is constantly applied, this results in stable limit-cycle oscillation. That is, even after a major perturbation the oscillator will quickly return to a particular path in phase space. A Fitzhugh-Nagumo oscillator is described by the following differential equation:

$$\begin{aligned}\frac{dv}{dt} &= -v(v - \theta)(v - 1) - w + \Omega \\ \frac{dw}{dt} &= \epsilon(v - \gamma w)\end{aligned}\tag{1}$$

where v is voltage, w is recovery of voltage, and Omega Ω is external driving voltage. In the simple case Ω is just a small amount of constant external voltage v_c necessary for sustained oscillation. However, when a rhythmical input pattern v_{in} is used, the driving term is $\Omega = v_c + v_{in}$. The frequency of oscillation is controlled by varying epsilon ϵ : a high value yields high-frequency oscillation with a sinus-like waveform while a low value yields low-frequency oscillation with a highly nonlinear waveform. This second, nonlinear oscillation is called *relaxation-mode* oscillation because the limit cycle of the oscillator closely follows or “relaxes onto” the cubic voltage v isocline of the oscillator. The two remaining parameters are fixed for all simulations. They include a firing threshold parameter θ set to 0.2, and a shunting parameter γ set to 1.2.

To simplify analysis, we modified the oscillator so that it generates a discrete output with every cycle. These outputs are generated each time the model neuron fires and are visible in Figure 1 in the flow diagram (left) and in the time series (center) of the oscillator.

We created a network by coupling the oscillators using a Heaviside coupling function. We chose this function based upon evidence [16] that Heaviside-coupled relaxation oscillators are exceptionally good at rapid and robust synchronization. As our goal is to synchronize quickly with important periodic components in the input, this was a desirable trait for the network to exhibit. The Heaviside function for a Fitzhugh-Nagumo oscillator $H(x) = 0.25(1 + \tanh((x - \tau)/\sigma)(v - 1.5))$ is in Figure 1, right. Though several values of τ and σ are shown in the diagram, these parameters were fixed for all simulations at $\tau = .7$ and $\sigma = .05$.

Flow diagram, time series, and coupling function

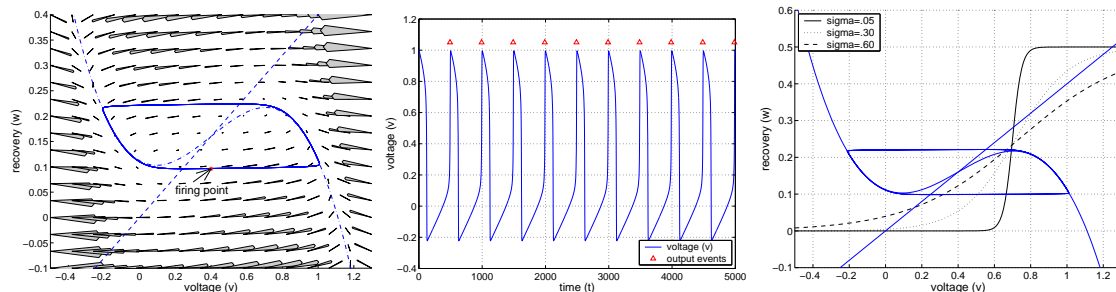


Fig. 1. The flow diagram (left) shows the the two distinct time-scales found in all relaxation oscillators: slow movement along the y axis and fast movement along the x axis. The time series (center) shows the oscillator waveform. In these plots, the firing point for output beats can be seen. The coupling function (right) is the Heaviside function with three values of τ and σ .

To add Heaviside oscillator-to-oscillator coupling, the driving term Ω in Equation 1 Equation 1 is modified to include voltage v_{net} from other oscillators: $\Omega = v_c + v_{in} - \alpha H(v_{net})$. Oscillator-to-oscillator coupling strength α is a variable that may be optimized. In fact, optimization through gradient descent learning is a current research goal. For these simulations, however, a weak positive connection $\alpha = .1$ is used for all oscillators.

3 Experiment 1: Finding Downbeats in A Set Of Rhythmical Patterns

To test the model we used a set of rhythms from Experiment 1 of Povel & Essens [15]. These rhythms are generated by permuting the interval sequence **1 1 1 1 1 2 2 3** terminated by the interval **4**. In this notation an integer indicates how much time follows a note. For example, an interval of 3 consists of 1 note followed by 2 rests. Thus all patterns are of length 16 and consist of nine notes and seven rests. These patterns were used by Povel & Essens in a learning experiment that revealed a connection between ease of downbeat induction and ease of pattern recall: though all patterns have the same number of notes and rests, patterns that are easy to tap along with were also easier for listeners to reproduce later. The authors propose a model of beat induction which predicts downbeat induction strength for temporal patterns. The patterns are assigned difficulty rankings from 1 to 7 based upon ease of downbeat induction (Table 1). In the figure, rests shown as . and notes as x.

3.1 Method

Input Patterns: The patterns in Table 1 were translated for the oscillator network by encoding notes as voltage spikes of width 1. So that the network of oscillators would run in low-frequency relaxation mode, the base interval between notes was 125 timesteps. Rests were not encoded in any way. The amplitude of the spikes started at .065 and increased evenly throughout the duration of the simulation to .08 (mean input strength .0725). This strategy of steadily increasing input amplitude kept the network from always preferring early notes in the input pattern. Early-note

1	xxxxx.xx.x.x...	xxx.x.xxx.xx...	x.xxx.xxx.xx...	x.x.xxxxx.xx...	x.xx.x.xxxxx...
2	xxx.xxx.xx.x...	x.xxxx.xx.xx...	xx.xxxxx.x.x...	xx.x.xxx.xxx...	x.xxx.xxxx.x...
3	xxx.xx.xx.xx...	xx.xxxx.x.xx...	xx.xx.xxxx.x...	xx.xx.xx.xxx...	x.xxx.xxx.xx...
4	xx.xxxx.xx.x...	xx.xxx.xxx.x...	xx.xxx.xx.xx...	xx.xx.xxx.x...	xx.xx.xxx.xx...
5	xxxxx.xx.x.x...	xxxx.x.xxx.x...	xxx.xx.xxx.x...	x.xxx.x.xxxx...	x.x.xxxx.xxx...
6	xxxx.x.x.xxx...	xx.xxx.x.xxx...	xx.x.xxx.xxx...	x.xxxx.x.xxx...	x.xxxxx.xx.x...
7	xxxx.xxx.x.x...	xxxx.xx.xx.x...	xx.xxxx.xx.x...	xx.x.xxxxx.x...	x.x.xxx.xxxx...

Table 1. 35 patterns ordered by row. In these patterns, an “x” indicates a note and a “.” indicates a rest. The difficulty ratings in the left-hand column are from a model by Povel & Essens and range from easy (1) to hard (7).

preference may seem like a weakness in our model, but in fact listeners exhibit the same preference [4].

Network Simulations: 20 Fitzhugh-Nagumo oscillators were instantiated at 4 times the base interval of the patterns, or period 500 ($\epsilon = 0.0015$). This choice of period simplified the task by giving the network the correct tapping rate. But it did not trivialize the problem because the network must still oscillate at the right phase with respect to the input. To achieve this, the oscillators must be stable enough to respond to some notes in the input (those indicating the downbeat) while ignoring all others. Oscillators were fully coupled to one another with a connection strength α of .01. This value was found by trial and error, and is similar to the value used in single-oscillator simulations [?]. At the start of simulation the 20 oscillators were distributed high on the left-hand side of the cubic voltage v isocline. Through simulation it was verified that the distribution was broad enough to ensure that the oscillators did not have an initial bias towards any particular downbeat. For each input pattern, the system was run for 8 pattern repetitions. All simulations were run using the ODE-solver suite in Matlab 5.3. Several solvers were tested to ensure that computational error did not affect results.

Measurement: Downbeat induction was measured using a binning method that matched oscillator outputs to locations in the input pattern. After allowing the system to run for 6 pattern repetitions, oscillator output for 2 following pattern repetitions were placed into one of four bins that represented the four possible downbeat phases. These pattern phases are shown at the bottom of Figure 4.

If the bins did not match for both pattern repetitions, an oscillator was declared “failed”. Thus, successful oscillators are those which always fired at the same phase and so revealed stable period oscillation. Through a large number of simulations, this method was shown to be an effective measure of beat induction. Namely, using more than 6 repetitions did not yield a significant decrease in failures. Also, using more than 2 binning repetitions did not yield in a significant increase in failures.

3.2 Results

Since downbeat induction is a perceptual phenomenon that changes with musical experience and presentation rate [11], there is no easy way to know in advance desired network behavior. That is, even with close examination it is not always clear where the downbeat *should* fall in a given pattern, especially for difficult patterns. Thus, it would be preferable to compare performance to

a large group of listeners performing beat induction on these same patterns. Unfortunately to our knowledge those experiments have not been done; they are the topic of future research. For this analysis we compared the network downbeat inductions to the predictions made by the Povel & Essens model mentioned above. This model was chosen due to its success at predicting ease of pattern learning in the same pattern set. A discussion of the model is not in the scope of this paper and readers are referred to [?] and [?] for our views on the model.

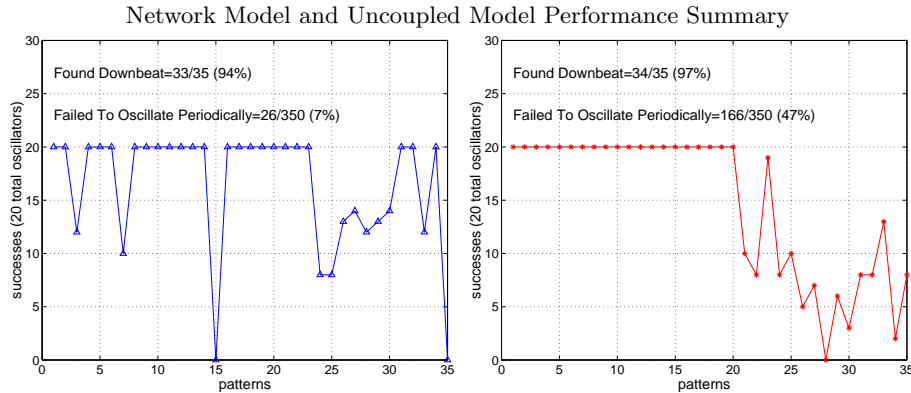


Fig. 2. Performance summary for the network model (left) and the uncoupled model (right).

In general the performance of the network was very good¹. Multiple oscillators locked onto the most strongly induced downbeat in 33 of 35 patterns (94%). In patterns 23 and 33 the Povel & Essens model predicts that two competing downbeats are equally likely. For both of these patterns, 12 oscillators discovered one downbeat while 8 oscillators discovered the other one. In Figure 2 (left) the number of successful oscillators is plotted for the patterns.

4 Discussion

In previous research we ran the same simulations with coupling turned off to test the usefulness of oscillator-to-oscillator coupling. These results are reported in detail in [?] and are summarized very briefly here.

As can be seen in Figure 2 (right), the uncoupled oscillator model did succeed at finding downbeats in many of the patterns. In fact, the uncoupled model failed completely at only one pattern while the coupled model failed at two of them. However, a closer look at the results shows clear advantages for the network model. In terms of failed oscillators—that is, oscillators that did not settle into a periodic limit cycle—the uncoupled model had 166 of 350 failures (47%) while the network model only had 26 of 350 failures (7%). This reveals a general lack of stability in the uncoupled model. Compare the performance summary of the coupled network model (Figure 2) (left) to that of the uncoupled model (right). Here it is easy to see that the uncoupled model suffers a drastic degradation in performance for patterns 21 through 35.

¹ The full set of predictions is available at <http://www.idsia.ch/~doug/icann2001data.html>

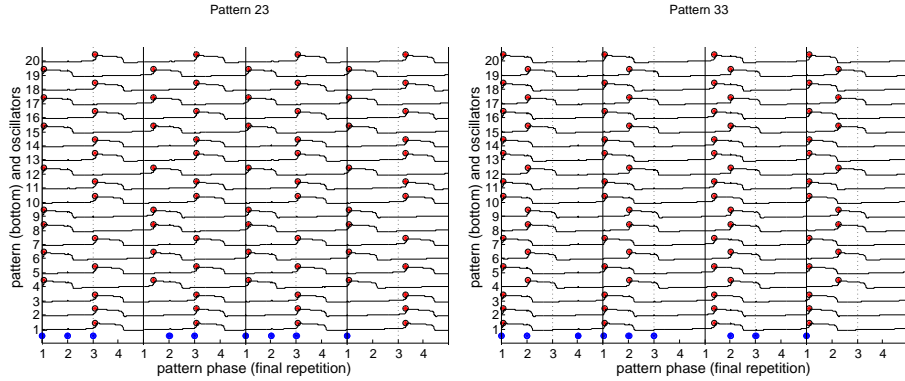


Fig. 3. In patterns 23 (left) and 33 (right), two downbeats are equally likely. The network finds them both.

In patterns 22 and 23 the Povel & Essens model predicts two equally-plausible downbeat phases. In these cases, the network found both downbeats. This can be seen in Figure 3. For pattern 23 (left) 8 oscillators are aligned with phase 1 and 12 with phase 3. For pattern 33 (right) 12 oscillators are aligned with phase 1 and 8 with phase 3. Furthermore, for patterns 17, 18 and 20 the Povel & Essens model predicts that phase 2 is a good downbeat assignment (though for reasonable parameter settings it is not as good as phase 1). For all three of these patterns, most oscillators are aligned with the preferred phase 1 while a small but significant number of oscillators are the second good phase 2. The uncoupled model does not exhibit this behavior. Though some oscillators are aligned with both solutions for patterns 23 and 33, in the case of patterns 17, 18 and 20 the uncoupled model only finds the primary solution of phase 1.

We believe that the behavior of the network is desirable. In complicated rhythmical patterns there is rarely only one important periodic component. The network model is superior to the uncoupled model at finding and synchronizing with multiple periodic components in parallel. This behavior, we believe, can be useful when more difficult tasks like pattern learning are considered.

The network failed most drastically on pattern 35. For this pattern, the network found two aperiodic solutions that show a sensitivity to grouping influences rather than metrical influences. Specifically, the two groups of oscillators lie close to periodic solutions at phases 1 and 3 but are pulled away due to influences from the In [15] it is suggested that for very difficult patterns like this one, a clock may not be induced at all and that listeners may use such non-metrical grouping structure to help memorize the pattern. [1] presents evidence that even young children are sensitive to both metrical and phrasal information in rhythms like these.

Two main observations have been made about the network of oscillators. First, it was shown that the network is less prone to drastic loss of efficiency for certain patterns than is the uncoupled model. Second, it was shown that the network finds multiple correct solutions more often than the uncoupled model.

These two phenomena can be attributed to *fractured synchrony* in the network model. Fractured synchrony [16] occurs when a group of coupled oscillators splits into small synchronized pools having stable phase relationships to one another. In networks with no external driving rhythmical pattern, the phase relationships between the pools are arbitrary. Here, with the presence of an

Two Examples of Failure

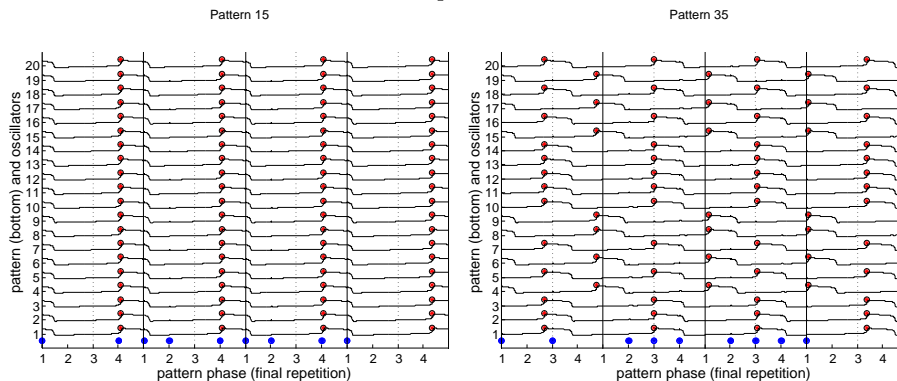


Fig. 4. On the left, (pattern 15) the network fails by finding downbeat at phase 4 instead of phase 15. On the right, (pattern 35) the network fails by oscillating aperiodically.

input pattern, multiple pools also emerge, but their phase relationships are no longer arbitrary but rather fixed by regularly recurring notes in the temporal pattern. A fractured synchrony explanation would predict that networks started with random initial phases should find multiple simultaneous solutions by pooling together (and staying together) due to the influence of positive oscillator-to-oscillator coupling; in addition, their phases should be fixed with respect to the input pattern due to oscillator-to-input coupling.

This explanation predicts that networks are more stable than single oscillators because oscillator pools are more difficult to perturb. This will yield less aperiodic oscillation in response to stray notes, thus explaining the network’s better performance on patterns 21 through 35. Also, more varied solutions should be found due to the fact that a pool of oscillators will be more likely to *remain* phase-aligned with some recurring energy in the signal once it has found that energy. That is, all pools will not collapse onto the *earliest* predicted solution.

In short, the positive oscillator-to-oscillator coupling acts as a substitute for signal energy when the signal does not provide a reliable note every time the oscillators reach phase zero (i.e. patterns 21 through 35). Uncoupled oscillators have no such mechanism. They either succeed at finding the earliest predicted solution when one is present or they oscillate aperiodically, causing them to be treated as failures by the binning mechanism.

5 Conclusions

These simulations successfully show that networks of Fitzhugh-Nagumo relaxation oscillators can perform beat induction. Given that these oscillators were designed to model the dynamics of neural action potential—not rhythm cognition—this is an interesting finding. This study is only preliminary, and clearly more work needs to be done. A comparison should be performed between this model and similar models (e.g. [8],[7],[12]). Experimental data needs to be collected from listeners performing a similar task. Finally, a pattern set which does a better job of finding the breaking points of this model and others like it would be of great value.

Though only 20 oscillators were used in these simulations, the dynamics of Fast Threshold Modulation predict that much larger networks should work without catastrophic desynchronization. Thus, the research in this chapter can be seen as a base from which to explore network models where a large number of oscillators can work in parallel to find important periodic components in temporal patterns. One obvious application of such a model is in learning long temporal patterns having rhythmical structure.

References

1. Jeanne Bamberger. *The Mind Behind the Musical Ear: How Children Develop Musical Intelligence*. Harvard University Press, 1991.
2. Grosvenor Cooper and Leonard B. Meyer. *The Rhythmic Structure of Music*. The University of Chicago Press, 1960.
3. R. Fitzhugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1:455–466, 1961.
4. W.R. Garner and R.L. Gottwald. The perception and learning of temporal patterns. *Quarterly Journal of Experimental Psychology*, 20:97–109, 1968.
5. Stephen Handel. *Listening: An introduction to the perception of auditory events*. MIT Press, Cambridge, Mass., 1989.
6. A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952.
7. Edward W. Large and Mari Reiss Jones. The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1):119–159, 1999.
8. Edward W. Large and J. F. Kolen. Resonance and the perception of musical meter. *Connection Science*, 6:177–208, 1994.
9. F. Lerdahl and R. Jackendoff. *A Generative Theory of Tonal Music*. MIT Press, Cambridge, Mass., 1983.
10. H.C. Longuet-Higgins and C.S Lee. The perception of musical rhythms. *Perception*, 11:115–128, 1982.
11. J. D. McAuley and Peter Semple. The effect of tempo and musical experience on perceived beat. *Australian Journal of Psychology*, 51(3):176–187, 1999.
12. J.D. McAuley. *On the Perception of Time as Phase: Toward an Adaptive-Oscillator Model of Rhythm*. PhD thesis, Indiana University, Bloomington, IN, 1995.
13. J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceeding IRE*, 50:2061–2070, 1962.
14. R. Parncutt. A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11:409–464, 1994.
15. D.J Povel and Peter Essens. Perception of temporal patterns. *Music Perception*, 2:411–440, 1985.
16. David Somers and Nancy Kopell. Waves and synchrony in networks of oscillators of relaxation and non-relaxation type. *Physica D*, 89:169–183, 1995.