

Learning and the Evolution of Melodic Complexity in Virtuoso Jazz Improvisation

Paul Hodgson (paulho@sussex.ac.uk)

COGS/Informatics, University of Sussex, Brighton BN1 9QH,
United Kingdom

Abstract

Spontaneous musical improvisation in jazz depends upon a substantial cognitive scaffolding that enables the real-time generation of creative structures. This work describes a computational model that simulates sub-conscious learning processes through an evolutionary paradigm.

Introduction

This paper focuses on the work of Charlie Parker and describes a computational model of evolutionary learning designed to simulate the process of dynamic phase construction in “off-line” and real-time improvisation.

In order to understand musical improvisation at the creative level of Charlie Parker we have to understand something about the generation of complexity. To take individual notes and combine them in ways that consistently produce creative music of the highest standard is to understand something about how simple elements are organised into complex structures. To sit a monkey at a piano and expect it to play a solo in the style of Parker is as amusing as it is improbable. To quote Johnson-Laird (1991):

“ There are many ways to break the rules of any genre: almost all of them are uninteresting and aesthetically unappealing. Geniuses need to know more, and to have this knowledge in a form that can control the generation of new ideas.”

Clearly anyone listening to a style of music over a prolonged period will absorb this stylistic knowledge at different levels. Research in cognitive musicology has focused on experiential learning Krumhansl (1995), and Jackendoff (1991) with the aim of understanding how this process takes place and how this information is used by listeners to guide their expectations and predict what will come next. Many simulations in music cognition incorporate implicitly or explicitly, certain a priori stylistic knowledge into their models. Lerdahl and Jackendoff's (1983) generative theory models what they describe as an “experienced listener”, and others (Rowe, 1992 and Witten & Conklin, 1995) pre-program a sense of tonal style into their systems. Reis (1999) developed a model to determine how far a system without any a priori stylistic information could learn about musical style by using context models.

Previous work on music learning has focused on building music grammars that have tended to be over complex and

rigid (Roads, 1996). Connectionist models have offered greater flexibility (Page, 1991) but have proved difficult to analyse. Context models (Witten et al, 1990) extract and accumulate a set of discrete sequences taken from the musical surface. Each pattern is then used as a context for the purpose of matching incoming data and generating expectations. Context models can learn with experience as new patterns are added to the database and exiting ones are reinforced by use of a frequency counter. These are then used as weights when multiple contexts are activated simultaneously. Reis implemented his model using perceptually guided segmentation, whereby data is segmented according to perceptual cues such as changes in registral direction or large jumps in pitch or time domains. Applying his model to 100 Bach chorales he found that prediction performance improved with experience but degraded with increasing prediction horizon.

There are a number of problems with this approach as far as creative musical improvisation goes. Firstly, unless a vast body of data is used as “learning experience” together with a complex probabilistic weighting system, the model will tend to produce very deterministic responses. Even if weights are used vast numbers of rules would have to be employed to create different and original phrases that remained faithful to the musical style, which could easily render the model computationally real-time intractable. The main feature of Parker's improvisation is the fact that it is not only stylistically consistent but is also creatively consistent. The ability to surprise and create in that way is unique to artistic genius.

A second problem for a statistical approach to musical style is that one can find that anything follows anything else if one studies enough data. As pointed out by Narmour

“...from an implicative point of view, we say that style does not exist as a unified, consistent whole. No sequence of events is a priori the seeming cause of another.” (Narmour, 1990 p25)

To try and understand the creative generative process in the work of Parker in greater depth I have however partly adopted this approach and returned to the source and taken a much more detailed look at solo construction in the original work. It is clear that Parker uses a motivic approach to improvisation but what is not clear is the nature of these motifs and the way in which they are altered and combined to create new phrases in real-time. Owens (1974 p17) work on Parker concludes that Parker's motives fall into several

categories. Some are only a few notes long and are adaptable to a wide variety of harmonic contexts and tend to be the most frequently used motives, occurring in virtually every key and piece. Others form complete phrases with well defined harmonic implications, and are correspondingly rare. Most motives occur on a variety of pitches, but some are confined to one or two pitch levels. A few occur only in a single group of pieces or a single key.

EarlyBird

EarlyBird is very similar to “Improviser”, “Virtual Bird” and “Gene-e” (Hodgson 1989, 1994, 1996) in that it sequences four tracks of music over MIDI to a sound module. EarlyBird was originally constructed as an evolutionary music game program along the lines of (Dawkins, 1986) Biomorph program. That is to say it contained a target phrase and random initialisation of eight “child” tracks from which the user selected two “parents” at each generation. It uses crossover and a variable user controlled mutation rate. Each track contains four bars of music and each bar contains sixteen squares on the x axis and 12 squares on the y axis. The user can input notes on the grid using different colours to represent different note values. The aim of the original program was to allow the user to evolve the target track, which can be heard but not seen, in the minimum number of generations.

The problem with this design from the perspective of understanding the generation of Parker’s music is that it evolves musical offspring towards **one** particular target phrase from which it is impossible to generate more general knowledge about Parker’s style. In doing this it evaluates on a note to note basis in the same way as Todd & Werner (1999) compare “bird song” notes using song transition tables. This says nothing about interrelations between notes on a more hierarchical level or in terms of more general musical structure. The key question in understanding Parker’s music is how to consistently generate solos that are in “Parker space” and not just particular instances.

Clearly Parker uses motifs but they are not developed or applied in a uniform way. In fact they depend very much on the melodic, harmonic and rhythmic context in which he is working. It is completely incorrect to suggest that Parker superimposed a standard set of motives to different situations. A close examination of his work reveals that phrases are being crafted to different situations in a way that must combine motives with a highly creative sensitivity to a particular situation. I decided to examine much more closely the identification of motives in Parker’s work. To do this I focused on one particular tune in Parker’s repertoire. The rationale for doing this was that firstly, and most importantly, there are twenty three known recordings of Parker playing this tune that I had access to. Secondly, I know and like the tune very much and have myself improvised on it extensively as an alto saxophonist for the last thirty years.

Parker’s favourite chord structure in the concert key of G major is the popular song “How high the moon”.



Figure 1: First two bars of How High the Moon by Morgan Lewis

As the alto saxophone is pitched in E flat it means any concert key has to be transposed down a minor third which places it in E major on the alto. In discussing this tune I will work in the transposed alto key. This tune is a thirty two bar tune consisting of two sixteen bar sections forming an “AA” structure. This tune was widely known by jazz musicians and played extensively throughout the 1940s and 1950s and was a tune used by Parker to compose one of his own original melody “Ornithology”, which is based on the harmonic changes of “How high the moon”. The title “Ornithology” refers to Parker’s nickname “Bird” and the first phrase consists of a classic upbeat Parker statement.



Figure 2: First two bars of Ornithology by Charles Parker

Owens (1974) identifies fourteen of his sixty four motives as being predominantly used on this tune. In relation to “Ornithology” he states that the approach to improvisation is the same on all the recordings he transcribes over the period because the tempo range is narrow (between 190 and 230 bpm). The takes are taken over a six and a half year period (3/46-9/52) during which Parker must have improvised over this tune countless times.

The structure of the first four bars consist of a two bar phrase in E major followed by a two bar phrase in E minor, “Somewhere there’s music, how faint the tune”. As the structure of the tune is AA these two bars are played twice per chorus. As there are fifteen choruses in the transcriptions there are a total of thirty instances where Parker improvises over the two bars of E major. I decided to focus in on these two bars and see exactly how Parker improvised over a small section of E major.

I could only find five motives that Owens identified. Admittedly, I have only looked at two bars in E major, whereas Owens applied his motives to the whole tune. However, this lead me to believe that Owens tended to identify motives that were too coarse-grained or fully formed. In fact a proportion of the motives Owens identifies are motives of one bar or more.

The 30 instances of the first two bars were transcribed and used as a data set in this research. What is astounding about these thirty instances is that almost all are completely different and all extremely melodic. This is an amazing feat given that most musicians tend to repeat their favourite lines

on tunes, and perfectly demonstrates why Parker was considered to be a creative genius. Furthermore, every note of the chromatic scale is used in the construction of these tonal phrases at some point. In fact if these two features were taken as a measure of creativity they could not be improved upon.

Having identified these 30 phrases in E major I now considered that I had the opportunity to use these phrases to try and evolve similar phrases. However, the main problem was how to extrapolate some kind of generality from this body of examples, as well as mixing new phrases from an existing population through crossover and mutation. This was the process employed in the “Virtual Bird” program which used ready made variable length phrases taken from transcriptions in exactly this way to explore the space. It is also clear, having reviewed the literature, that others who have tried to generate Parker structure using other techniques such as neural nets and genetic algorithms have not had much success as described in Spector & Alpern (1995) and Biles, Anderson & Loggi (1996).

Taking all thirty examples of Parker improvisations over the first two bars of Ornithology as a data set of motives similar in size to those of Owens, I used EarlyBird to “evolve” new structures using crossover and mutation. The results however, were variable and in no way exhibited the fluidity and coherence of the original phrases. I encountered the same problem as I had with the Virtual Bird program, which uses a database of Parker phrases but takes them out of context and applies them in a completely mechanical way.

Parker had to learn how to play such melodic music and his musical style is documented as having developed rapidly over a relatively short period of time (Owens, 1995). Instead of using statistical techniques to try and extrapolate structure from musical examples to build complex syntax, I decided to reduce the level of granularity to the most basic level I could beyond the single note level. Having done this I would then be in a position to generate larger structures from the most primitive, in a way that would maximise the space of possibilities and hopefully retain the musical style. I reduced the data set of thirty two bar E major examples down to a set of some fifty primitive structures. Most of these are dyads (two note structures) as I wanted to use the simplest possible structures beyond the individual note level. These structures are key to this simulation as they effectively allow the generation of a set of fitness peaks in a fairly large landscape of 31³².



Figure 3: Example dyadic primitives

Narmour (1990, 1992) posits two inter-dependent perceptual systems, bottom-up and top-down. In the bottom-up system pitch intervals vary in the degree of closure they convey. An unclosed interval is considered implicative and generates expectancies for the following

realised interval. Narmour considers dyads as unrealised implications of registral direction and intervallic motion. They contain initial and terminal tones, either of which creates the transformational structural tone. Narmour considers that “Dyads are full-fledged melodic structures, and, as such, they can occur contiguously in a series...” (Narmour, 1990 p 398) and he also considers that dyads can dominate process if the phrasing and context are unusual. He considers that process tends to dominate dyadic structuring because in process similarity is strongly cohesive, and intervallic and registral sequences of similarity resist disjunction. Qualifying this he states “Thus it follows that process lying at the defined threshold of similarity (i.e. close to differentiation) can become dyadic if phrasing and context make it so.”(ibid., p401)

If this is the case I reasoned that improvisation could be seen as a subconscious process of reordering implicative primitives into larger implicative/realisational structures. Dyads are implicative and give the largest possible space of possibilities beyond individual notes or monads.

Melodic and rhythmic information is stored separately in the program and the pitch information is stored as pitch classes. The aim now was to connect these primitives together to form new two bar phrases that were entirely consistent with Parker’s style.

The idea that for any particular tune, Parker stored a mental map of primitive dyadic structures that could be recombined rapidly, is an interesting idea and it needs to be tested. It could be the case that these structures are not only different for different tunes but also for different sections of tunes. The implication is that a huge part of the evaluation is done pre-improvisation. Jazz musicians spend a vast amount of time “working over the changes” to particular tunes. This process is akin to a sculptor paring down a block of wood to the rough shape of the desired object, or the way in which theoreticians prune irrelevant ideas. If the improvising musician already has a set of structures in memory that can be combined and modified very rapidly in real-time, this would more easily explain how complex structures can be built spontaneously.

To test this idea I started by using random numbers to chain structures together. Instead of relying on the pseudo random number generator on the computer to generate pseudo randomness I decided to use lottery numbers to generate true randomness. From listening to these it was evident that these phrases were awkward and not what one would consider, flowing melodic lines. To remedy this situation I decided to impose some structure from without and insisted that the pattern chaining process had to follow one of the melodic contours (ascending, descending) from the original phrases. The set of primitives was divided into an ascending, descending and lateral set and each primitive was scored for the number of repetitions. Random choices were then selected from only those primitives that were repeated one or more times. These were also considered unsatisfactory because there was too much repetition for

“It is, after all, reasonable to imagine the existence of a range of listeners who would invoke a different set of similarity/difference rules from those I have advocated in the implication-realisation model....what one hears, after all, depends on how one perceives similarity and differentiation.” (Narmour, 1992 p359)

Acknowledgments

I would like to thank Margaret Boden for giving me the opportunity to pursue this work, and for providing creative supervision.

References

- Biles, J.A., Anderson, P.G., Loggi, L.W. (1996). Neural Network Fitness Functions for a Musical IGA. Technical Report, Rochester Institute of Technology.
- Boden, M. (1993). *The Creative Mind*, Basic Books.
- Dawkins, R. (1986). *The Blind Watchmaker*. Norton, New York.
- Hodgson, P.W. (1989) *Understanding Computers, Cognition and Creativity*. MSc Thesis, University of West of England.
- Hodgson, P.W. (1989). *The Improviser Jazz Tuition System*, published by Creative Sounds Ltd, Bristol, UK.
- Hodgson, P.W. (1994). *Virtual Bird jazz simulation system*.
- Hodgson, P.W. (1995). *Modeling Cognition in Creative Musical Improvisation*. In D.Wessel, (Ed), *Proceedings of the Society for Music Perception and Cognition*. University of California, Berkeley.
- Hodgson, P.W. (1996). *Gene-e evolutionary Dance Music system*.
- Hodgson, P.W. (1999). *Modelling Cognition in Musical Improvisation through Evolution*. *Proceedings of the AISB99 Symposium on Musical Creativity*. University of Edinburgh.
- Hodgson, P.W. (2002). *Artificial Evolution, music and methodology*. *Proceedings of the 7th International Conference on Music Perception and Cognition*, Sydney.
- Hodgson, P.W. (2005). *Modelling cognition in creative musical improvisation*. Dphil Thesis, University of Sussex.
- Jackendoff, R. (1991). *Musical Parsing and Musical Affect*. *Music Perception*, Vol 9, 199 – 230.
- Johnson-Laird, P. N. (1991). *Jazz Improvisation: A Theory at the Computational Level*. In *Representing Musical Structure*, eds. P. Howell, R. West, and I. Cross. San Diego, Calif.: Academic.
- Krumhansl, C.L. (1995). *Music psychology and music theory: problems and prospects*. *Music Theory Spectrum*, 17: 53-80.
- Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*, MIT Press.
- Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures*, University of Chicago Press.
- Narmour, E. (1992). *The Analysis and Cognition of Melodic Complexity*, University of Chicago Press.
- Owens, T. (1974) PhD Thesis on Music of Charlie Parker. University of California.
- Owens, T. (1995) *Bebop the music and the players*. Oxford University Press.
- Page, M. (1991). *The musical expectations of self-organising neural networks*. In *Proceedings of the international computer music conference*, 206-209.
- Reis, B.Y. (1999). *Simulating music learning: On-line, perceptually guided pattern induction of context models for multiple horizon prediction of melodies*. *Proceedings of the AISB99 Symposium on Musical Creativity*. University of Edinburgh.
- Roads, C. (1996). *Computer Music Tutorial*. MIT Press.
- Rowe, R. (1992). *Machine listening and composing with Cypher*. *Computer Music Journal*, 16(1):43-63.
- Spector, L. & Alpern, A. (1995). *Induction and Recapitulation of Deep Musical Structure*. In *Proceedings of the IJCAI-95 Workshop on Artificial Intelligence and Music*.
- Todd, P.M. & Werner, G.M. (1999). *Frankensteinian Methods for Music Composition*. In *Todd, P.M. & Griffith (Eds.), Musical Networks and Performance*. MIT Press.
- Witten, I.H., Bell, T.C & Cleary, J.G. (1990). *Text Compression*. Prentice Hall, Eaglewood Cliffs, NJ.
- Witten, I.H. & Conklin, D. (1995). *Multiple viewpoint systems for music prediction*. *Journal of new music research*, 24:51-73.