An aesthetic comparison of rule-based and genetic algorithms for generating melodies

ANDREW R. BROWN

Queensland University of Technology, Victoria Park Road, Kelvin Grove 4059, Brisbane, Australia
E-mail: a.brown@qut.edu.au

Various algorithmic techniques are available for generating music, many of which come from the field of artificial intelligence, which is rich with potential in this regard. However, the musical appropriateness of these techniques is less clearly understood. In this paper, I will report on a study that aimed to describe the characteristics of two of these techniques, rule-based and genetic algorithms, as they apply to melody generation. The appropriateness of these characteristics in contributing to well-formed melodies was judged by aesthetic criteria. The results indicate that most combinations of rules, mutations and evolutionary selection result in poor or average melodies, but that careful combination of these techniques can generate melodies that are not simply well-formed but in many cases display some elegance and novelty.

1. INTRODUCTION

> When norms are violated *ignorantly* (i.e., by a novice), the results are likely to be bad, whereas when they are violated *knowingly* (i.e., by an expert), the result is likely to be just fine. (Hofstadter 1995: 429)

Understanding the creative process is an ongoing quest for many researchers, including Douglas Hofstadter, and it is clear that various algorithmic processes mimic aspects of this creative process to varying degrees. This paper will describe the characteristics of two such processes as they were applied to the generation of musical melodies. The two processes are rule-based and genetic algorithms. The questions of interest for this study include: What are the musical tendencies of various algorithmic methods? and; How might these processes be combined such that their characteristics complement one another?

A broad way of describing these two processes might be to depict them as corresponding to different types of creative expertise, with rule-based systems containing heuristics as used by experienced composers, and evolutionary systems as relying on 'blind' chance (Dawkins 1986), as might be expected of a novice composer. Characterised in this way it was interesting to notice that Hofstadter's observation of the difference between novice and expert creators was evident in the characteristics of these algorithms.

This paper will elaborate first on the central themes of aesthetic assessment and rule-based and genetic algorithm processes. There are two sources of generative algorithms' strategies that are explored in this paper. Firstly, rule-based algorithms, the applications of which are often described as knowledge-based or expert systems. Here rules are used as both constructors and constraints to build and limit melodic material. Secondly, genetic algorithm processes are explored as a means of providing variation through mutation and selection using a fitness function. An overview of the rules and genetic processes is discussed below. Each rule, mutation or fitness function acts in isolation, but their character is assessed by their impact on the overall form of the melody, in particular as it affects their aesthetic quality.

This paper will then outline the experiments that systematically explored combinations of those processes in generating melodies, and discuss the results of those trials. Finally a deliberately selected combination of rule and genetic processes is tested as a method of balancing the tendencies of each algorithm. The paper concludes with a summary of the findings and indications for further research.

2. BACKGROUND

Music has long been a useful context for artificial intelligence researchers to study algorithmic models of thought and action (Minsky 1981; Winograd 1986). For Marvin Minsky, musical structures provided insight into the way the mind works, and can be useful metaphors for understanding, temporal perception and knowledge. For Terry Winograd, the domain of music has been a useful context in which to explore notions of intelligence and understanding. For these researchers, music is abstract enough to absorb many otherwise complex theoretical structures, while sonification has made these structures accessible enough to easily observe the results of their simulations and experiments. The application of generative computational processes to specifically artistic ends is more recently emerging as a field in its own right, quite apart from artificial intelligence research (Dorin 2001). The research reported in this paper continues this latter trend, but takes one further step to acknowledge that music is not simply an experimental context, but that a

Organised Sound **9**(2): 193–199 © 2004 Cambridge University Press. Printed in the United Kingdom. DOI: 10.1017/S1355771804000000

musical understanding is a mode of aesthetic understanding, which is itself a valid and useful measuring stick of algorithmic success.

Choices based on aesthetics have been acknowledged as contributing to new knowledge in many fields (Boden 1990), and as being 'more important in computing than anywhere else in technology' (Gelertner 1998: 20). This importance comes from the fact that, according to David Gelertner, aesthetic judgements 'are our most reliable guide to achieving software's ultimate goal: to *break free of the computer*, to break free *conceptually*' (*ibid.*: 20, italics in original). The use of aesthetic judgements in this study to make assessments about algorithmic character positions aesthetics as a lens through which to analyse the algorithms. This lens provides a perspective that looks beyond efficiency and optimisation.

The significance of this research for generative music-making lies in the acknowledgement that algorithmic strategies have stylistic tendencies. As we develop a better understanding of these tendencies, the appropriateness of particular processes to specific compositional ends will become more clear. A search for an understanding of the musical tendencies of algorithms is a somewhat different investigation than the application of different algorithmic processes to the generation of different styles. However, the work in such application, which is often more complex in nature than the melody generation reported here, has informed the choice of algorithmic strategies and their application in this work (see, in particular, the research into melodic generation and harmonisation by Biles 1994; Jacob 1995; Horner 1991; Cope 1998; Wiggins 1998; Todd and Werner 1999; Santos *et al.* 2000; Johnson 2003).

3. MUSICAL AESTHETICS

Aesthetic judgement is viewed, for the purposes of this research, as a heuristic used to evaluate the appropriateness of rule-based and genetic algorithms. It is as a synthesis technique to arrive at a final decision about the value of particular combinations of generative processes. So, while the development of the processes involved a significant amount of empirical work, the final judgement is an individual and artistic one. This summarising feature of aesthetic judgement is upheld by Robert Duisberg, who states that 'the very act of knowing involves the ordering of sensory input according to conceptual frameworks, but the development and application of appropriate concept networks is necessarily a creative and intuitive act' (Duisberg 1984: 205). In searching for aesthetic value in the melodies produced by rules and genetic algorithms, I am looking for glimpses of elegance and novelty, and looking for a minimisation of incoherence and tedium. I am not looking for (nor claiming the possibility of) aesthetic criteria such as insight, affect or emotional expression. The role of affect in algorithmic music becomes a metaphysical debate beyond the scope of this paper, and has been discussed elsewhere (Duisberg 1984). However, influences of these traits cannot be excluded from any aesthetic judgement even though they were not the primary criteria.

The interpretive framework used to make the aesthetic judgements in this study is that of classical diatonic harmony or, more correctly, a caricature of it. Music of this genre was the source from which the melody-writing rules that underpin the rule-based algorithms were elicited, and it is a context broadly enough known that readers can make intuitive authentications of the results based on the limited descriptions and examples provided. In particular, the melodies were assessed for:

- tonal and rhythmic coherence,
- stability of melodic contour, and
- balance of repetition and variety.

That these aspects of melodic style are significant is reinforced by their prevalence in educational literature on melody writing and in observations of children's melody perception (Dowling 1988).

It should be noted that in this research there were deliberate restrictions placed on melodic sophistication to reduce the problem's complexity and so that the influence of the algorithms would be more easily made evident.

4. RULE-BASED PROCESSES

The rules for melody generation used in this research were, as mentioned, derived from texts designed to assist beginner composers. The rules concerned pitch selection, rhythmic construction, repetition, and melodic contour. Tests were undertaken to identify the degree to which these rules were reflected by a statistical analysis of melodies in the cannon of Western repertoire (Towsey, Brown, Diederich and Wright 2000). The most reliable rules were identified, which included rules relating to the following melodic features:

- *Pitch Variety* – the diversity of unique pitches.
- *Tonal Deviation* – the extent of non-scale pitches.
- *Dissonance* – the sum of the dissonance rating of each note where each pitch set has a particular rating.
- *Overall Pitch Direction* – the balance of ascending or descending pitch steps.
- *Note and Rest Density* – the average number of notes or rests per beat.
- *Rhythmic Variety* – the diversity of unique rhythm values.
- *Syncopation* – the extent to which notes are held across bar or beat boundaries.

- *Repeated Pitch Density* – the extent of pitch repetition.
- *Repeated Pitch Patterns* – the extent of repeated intervallic patterns.

Rule-based systems have been used in computing for a long time and are the basis for knowledge-based or 'expert' systems used in artificial intelligence. Their behaviour is well understood. In rule-based systems, the encoded knowledge is explicit. The systems behave predictably, which contributes to their utility, but they are not necessarily deterministic. However, they do tend to be brittle, that is, they break down easily when conditions or context is altered.

5. GENETIC ALGORITHM PROCESSES

Genetic algorithm (GA) processes are inspired by evolutionary biology, in particular processes of generational breeding – combining elements from successful outcomes to produce new, hopefully better, ones – and mutation – making changes in an attempt to improve design. GAs are frequently used to search for better (more optimal) solutions to a problem. This search space is often described as an uneven landscape, where height corresponds to solution quality such that hill-climbing is the goal. Variations introduced through mutation of the melodies aim to help the individual melody take a step up the fitness-score hill, thus indicating a better solution. Keeping to this metaphor, it should be noted that the landscape of musical melodies is extremely uneven and the location of the highest mountains is unpredictable. Given this difficult terrain, it would appear that GAs are not likely to be an effective tool for locating well-formed melodies.

However, GAs offer advantages beyond optimisation. For example, when simulating human performance they preserve a trade-off between novelty and structure. It is this feature of GAs that was exploited in this study. The process involved 'breeding' populations of melodies with individuals undergoing 'mutations' followed by a selection of the 'fittest' melodies for use as a basis for the next 'generation'. GAs were applied to melody improvement not primarily because of their tendency to find *optimal* solutions, but because of their ability to generate *variety* through mutation; not because they were good mountain climbers but because they could find interesting mountains to climb.

An issue when implementing GAs for music is how the music will be encoded as a 'gene' so that it can be bred and mutated. In this study, a musical note was considered to be a gene and a phrase to be a gene strand. Note objects contained the required parameters of pitch and duration, and phrases containing an ordered series of notes.

Two methods of generating an initial population were explored. One was to begin with melodies made up of randomised pitch and rhythmic values, the other to begin with melodies generated by rules. When breeding a new population, two types of mutations were explored: conventional evolutionary techniques and deliberate 'musical' variations. There were two evolutionary mutation processes employed:

(1) *Random Pitch* – change one note's pitch in the melody each generation.
(2) *Split and Merge* – divide the melody at beat boundaries and recombine sections with those from other melodies.

Musical mutations were developed based on compositional rules of variation. The probability that each rule is applied was set at twenty per cent for all experiments.

(1) *Bar Sequence Mutations* – repeat a short portion of the melody with a small diatonic transposition.
(2) *Step Interpolation* – insert a note in the middle of a large intervallic leap and halve the note value of the first and inserted note to preserve beat locations.
(3) *Tonal Pauses* – lengthen a triadic tone on a downbeat by tying it to the following note.

The 'best' melodies are selected from the population based upon a fitness function. Different types of fitness function, or critic, have been explored for GAs; most common are the human critic and the rule-based critic (Towsey *et al*. 2000). In this study, another approach to the problem of constructing a fitness function was used, incorporating statistics obtained from the analysis of a library of melodies. The significant melodic features listed above were also employed in the fitness function. Means and standard deviation from a test population of several hundred melodies from the classical canon were derived. Generated melodies whose feature scores were more than two standard deviations from the mean would receive a low fitness rating. The two highest-scoring melodies in each population were maintained, and a new population was created by generating and mutating additional melodies.

6. EXPERIMENTS

A systematic test was carried out to assess which elements of the rule and GA processes contributed positively to the generation of well-formed and interesting melodies. Attention was paid to the characteristic contribution of each process to the aesthetic outcome. Finally, a selected combination of the processes that balanced the characteristics was tested. At each stage many listening trials were undertaken. Representative samples are presented in common-practice notation throughout the discussion.

Table. An overview of the results of assessment by aesthetic judgement.

Beginning state	Unprocessed	Fittest unmodified	Evolutionary mutations	Musical mutations	Combined mutations
Random	☆	☆	☆	☆☆	☆
Rule-based	☆☆☆	☆☆☆	☆☆	☆☆☆☆	☆☆☆☆

The melodies in this study were generated by a specifically written software package developed using the jMusic libraries in the computer language Java (Sorensen and Brown 2000). The generated melodies were saved as MIDI files, which were assessed by experienced musicians listening to synthesised performances of the files and by viewing scores of the melodies in common-practice notation.

The experiments in melody generation involved various combinations of rule-based and genetic algorithm processes. Two beginning states were utilised: randomly-generated and rule-generated melodies. Melodies resulting from these states alone were described as unprocessed. Genetic algorithms with selection and/or mutation processes were applied to melodies from each of these beginning states. The table summarises the combinations explored and displays a crude scoring of the aesthetic quality of the resulting melodies.

6.1. Random

The melodies is this group consisted of sixteen notes with randomly selected chromatic pitches in a two-octave range above middle C, and with randomly selected rhythmic values of either a quaver, crotchet, minim or semibreve. The result was, not unexpectedly, unorganised and quite unlike classical diatonic melody.

6.2. Random fittest unmodified

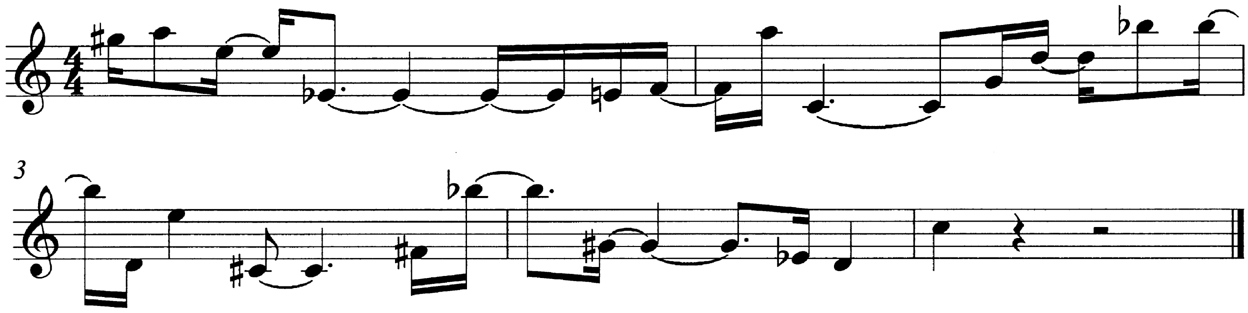
The melodies in this group were randomly generated, as above, but twenty generations of a population with fifty melodies were fitness tested. No changes or mutations were applied. The fitness scores were quite low but did increase during the first ten to twelve generations before reaching a plateau, after which there were occasional increases. The resulting melodies were similar to the random ones, with some greater tendency for diatonic pitches and less jagged melodic contours. This test indicated that the fitness selection process by itself was effective but not sufficient to direct the evolution when the candidates were so consistently poor.

6.3. Random evolutionary mutations

This stage was similar to the previous process with the addition of a random pitch and crossover mutation applied to each melody, at each generation. The results were indistinguishable from the random group. It appears that any positive tendency of the GA selection process was counteracted by the disrupting processes of the evolutionary mutations. In nature, these mutations provide variations on an already well-formed species, whereas in this case the any advantage of the variation process was lost amid the 'noise' of the randomised population. The minimal effects on a random starting population are not surprising given the limited number of generational iterations used in these experiments.

6.4. Random musical mutations

The melodies in this stage were varied by the musical mutations rather than by the evolutionary mutations. The musical mutations are effectively rule-based processes, designed to elaborate or to add stability to the melodies. The results at this stage showed clear evidence of diatonic and rhythmic order; however, this was very sporadic. The results of this stage clearly demonstrate that the purposeful variations are effective in pushing the melodies in the desired



1

Figure 1. Random melody.

direction. The fitness scores indicate that most of the progress occurred in the first few generations with the improvements slowing down considerably after that. It is likely that achieving melodies free of awkward moments with this process would take hundreds of generations.

6.5. Random combined mutations

In this group both evolutionary and musical mutations were combined and applied to a random starting population. The results were more diatonic and organised than the random starting melodies but displayed less frequent patches of coherence than when the musical mutations alone were used.

6.6. Rule-based

The melodies in this group were constructed by the rule-based system. There was no mutation or GA selection applied. The musical quality of these, taking into account the desired target style, was quite acceptable but very conservative. They were occasionally convincing but rarely novel.

6.7. Rule-based fittest unmodified

At this stage the rule-based melodies were filtered by the fitness selection. As with the random melodies, twenty generations of a population of fifty were tested for fitness. The best two melodies were maintained into the next generation. The melodies from this process were more consistently coherent than the previous stage, but the results were no more novel. The rule-based population initialisation meant that the GA search of the musical terrain started at reasonably high points, and the fitness process did not improve the results but was effective in eliminating the weaker candidates. Interestingly, at this stage the maximum fitness score reached a plateau after only five or six generations and rarely advanced beyond that in subsequent generations.

6.8. Rule-based evolutionary mutations

Until this stage, results of this study have generally confirmed conventional wisdom regarding the characteristics of rule-based and GA processes. If this trend was followed, the addition of evolutionary mutations (random pitch changes and structural crossover) would be the key to unlocking fitness progress by introducing variations, and noticeable improvements in the outcomes should be observed. However, while the results from these tests showed that the mutations added some interest to the melodies, this was usually in a quirky way that was generally unpleasant. Thus the evolutionary mutations had the opposite effect from what might have been predicted.

One explanation for this may lie in the encoding process where each note is a gene in a virtual DNA strand. Despite this being the most common type of GA encoding, it could be problematic because the genetic metaphor is not strictly adhered to. The change in a gene or a gene sequence in biological evolution does not entail a direct change in a species feature, such as dropping off an arm or adding a new organ. Rather the changes are indirect, affecting evolutionary processes rather than the melodies directly. An equivalent of this indirect connection for melody generation may be to encode the algorithmic parameters in the gene and affect them, rather than the notes directly.

6.9. Rule-based musical mutations

For this group, the rule-based melodies had musical mutations applied at each generation. There was a twenty per cent probability that each mutation would be applied. The results showed that the overall form returned to the conservative nature of the original rule-based melodies. The melodies demonstrated more variety than the unprocessed group and occasional flashes of novelty. The sequence and rhythmic subdivision mutations were particularly effective in this regard. Melodies in this group were the most aesthetically pleasing of any of the combinations.

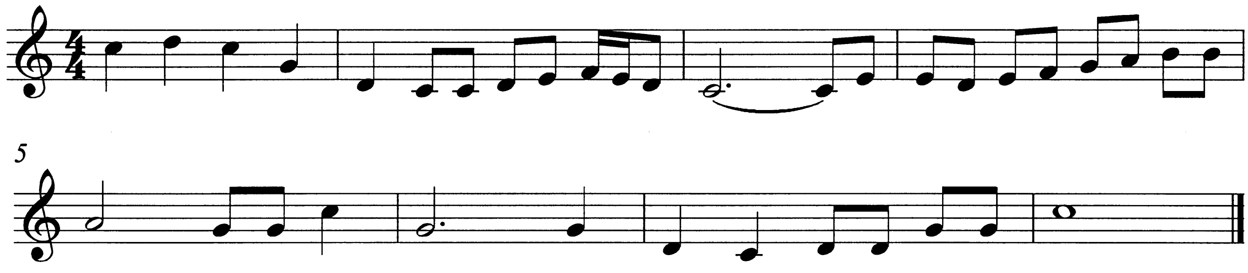
6.10. Rule-based combined mutations

Adding both types of mutations to rule-based populations produced melodies with significant variety, to the point where the melodies sounded fragmented or

Figure 2. Rule-based unprocessed melody.

1

1 **Figure 3.** Rule-based evolutionary mutation melody.



1 **Figure 4.** Best mix melody.

interrupted. Generally, the results were unsatisfactory overall, and whereas the simple rule-based originals were conservative, the combined mutation group were eccentric.

7. BEST MIX?

Taking into account the results of the experiments, a preliminary attempt at combining the rule and GA processes to maximise aesthetic effect was undertaken. Unsurprisingly, the process used rule-based generation to provide a coherent starting point. A combination of evolutionary and musical mutations was arrived at. The probability of each musical mutation was increased above the level used in the tests, with the Tonal Pause mutation increased to introduce a greater sense of phrasing which was often smeared by the other mutations. The random pitch change was eliminated altogether because the likelihood of a useful chromatic pitch introduction was extremely low. Here, in summary, are the final settings for the mutation levels:

- *Random Pitch* – 0 changes per generation
- *Split and Merge* – 1 change per generation
- *Bar Sequence Mutations* – 40% probability
- *Step Interpolation* – 40% probability
- *Tonal Pauses* – 60% probability

Finally, the GA breeding and fitness selection processes were employed. This eliminated many weak melodies and did little harm to stronger ones. The evolutionary process was limited to five generations, which captured most of the fitness increase while allowing for a relatively speedy generation process.

8. CONCLUSION

Aesthetic judgement is a useful device for assessing the characteristics of the melody generation potential of various algorithmic processes because it brings together a collage of rational, intuitive and contextual considerations. This is possible because 'affect and aesthetics are the very basis for knowledge, even purely factual knowledge' (Duisberg 1984: 231). It is also a pragmatic approach to the extent that a general audience for the generated material will focus mainly on its affect rather than the architectural details of its construction. This approach acknowledges that we understand music, as with other arts, through our experience of them. This view was asserted by John Dewey (1958) and considered at length in relation to musical aesthetics by Richard Shusterman (1992).

Some limitations of this study include the reliance on expressionless performances. There may be hidden potential aesthetic value in the melodies that may be extractable by human interpretation. Another limitation in the study may be that the rules for generation, mutation and fitness were closely related, and therefore mutually reinforcing, rather than setting up a creative tension between change and selection stages. However, this is nowhere near certain because even when such a tension was established through the use of random mutation, the results were poor.

It is reasonable to wonder if other algorithmic techniques may have been more effective. For example, other likely algorithmic candidates include Markov models, recurrent neural networks, and augmented transition networks (Cope 1992, 1998). The application of these techniques to music composition may well benefit from a similar aesthetic analysis of character. Future research may also explore the limitation exposed in the GA encoding model, by encoding generative parameters rather than notes in the genome to achieve second-order, perhaps more subtle and complex, influences.

The results of this study show that rule-based procedures produce well-formed melodies, but that novelty is rare. Limited musical mutations of these melodies can add interest and surprise. Evolutionary mutations, in particular random changes, tend to reduce coherence of form with minimal gain in elegance or novelty. These findings tend to support Hofstadter's assertion that designed (expert) deviations are more useful than stochastic (ignorant) ones in creative realms.

In the final analysis, the differences in algorithmic character were not so much between rule-based and evolutionary processes, as they were between random and deliberate construction and elaboration. Deliberate rules or mutations alone were too often unsurprising, while random generation or change usually led to inappropriate results. The outcomes of either process benefited by checks imposed by fitness selection that constrained inappropriate mutation and that filtered out unsatisfactory results; however, these checks were not always reliable. The choice for the generative system designer is to balance a reliable but conservative system on the one hand with an inconsistent but potentially interesting system on the other.

REFERENCES

Biles, J. A. 1994. GenJam: A genetic algorithm for generating jazz solos. In *Proc. of the Int. Computer Music Association*, pp. 131–7. San Francisco: International Computer Music Association.

Boden, M. 1990. *The Creative Mind*. London: Cardinal.

Cope, D. 1992. Computer modelling of musical intelligence in EMI. *Computer Music Journal* **16**(2): 69–83.

Cope, D. 1998. Signatures and Earmarks: computer recognition of patterns in music. In W. B. Hewlett and E. Selfridge-Field (eds.) *Melodic Similarity: Concepts, procedures, and applications*, pp. 129–38. Cambridge, MA: MIT Press.

Dawkins, R. 1986. *The Blind Watchmaker*. Harlow: Longman Scientific & Technical.

Dewey, J. 1958. *Art as Experience*. New York: Putnams.

Dorrin, A. 2001. Generative processes and the electronic arts. *Organised Sound* **6**(1): 47–53.

Dowling, J. 1988. Tonal structure and children's early learning of music. In J. Sloboda (ed.) *Generative Processes in Music: the Psychology of Performance, Improvisation, and Composition*, pp. 113–28. Oxford: Clarendon Press.

Duisberg, R. 1984. On the role of affect in artificial intelligence and music. In J. Rahn (ed.) *Perspectives on Musical Aesthetics*, pp. 177–93. New York: W. W. Norton.

Gelertner, D. 1998. *The Aesthetics of Computing*. London: Phoenix.

Hofstadter, D. 1995. *Fluid Concepts and Creative Analogies*. New York: basic Books.

Horner, A., and Goldberg, D. E. 1991. Genetic algorithms and computer assisted music composition. In *Proc. of the Int. Computer Music Conf.*, pp. 479–82. San Francisco: International Computer Music Association.

Johnson, M. D. 2003. *Evolutionary Computation Applied to Melody Generation*. web.umr.edu/~tauritzd/courses/cs401/fs2003/project/Johnson.pdf

Minsky, M. 1981. Music, mind, and meaning. *Computer Music Journal* **5**(3): 28–44.

Santos, A., Arcay, B., Dorado, J., and Rodriguez, J. 2000. Evolutionary computation systems for musical composition. In *Proc. of Acoustics and Music: Theory and Applications (AMTA 2000)*, vol. 1, pp. 97–102.

Shusterman, R. 1992. *Pragmatist Aesthetics: Living Beauty, Rethinking Art*. Oxford: Blackwell.

Sorensen, A., and Brown, A. R. 2000. Introducing jMusic. In A. Brown and R. Wilding (eds.) *Proc. of the Australasian Computer Music Conf.*, pp. 68–76. Brisbane: ACMA.

Todd, P. M., and Werner, G. M. 1999. Frankensteinian methods for evolutionary music composition. In Griffith and Todd (eds.) *Musical Networks: Parallel Perception and Performance*, pp. 313–39.

Towsey, M., Brown, A., Wright, S., and Diederich, J. 2000. Towards melodic extension using genetic algorithms. In A. Brown and R. Wilding (eds.) *Proc. of the Australasian Computer Music Conf. 2000*, pp. 85–91. Brisbane: ACMA.

Winograd, T. 1968. Linguistics and the computer analysis of tonal harmony. *Journal of Music Theory* **12**(1): 2–50.

Author Query Sheet

Manuscript Information
Journal Acronym JOS
Volume and issue 29:2
Author name Andrew R. Brown
Manuscript No.

QUERY NO.	QUERY DETAILS
1	Figure 1, 2, 3, 4, weren't mentioned in this paper