GENERATING MUSICAL ACCOMPANIMENT THROUGH FUNCTIONAL SCAFFOLDING

Amy K. Hoover
Dept. of EECS
University of Central Florida
Orlando, FL 32816-2362 USA
ahoover@eecs.ucf.edu

Paul A. Szerlip
Dept. of EECS
University of Central Florida
Orlando, FL 32816-2362 USA
paul.szerlip@gmail.com

Kenneth O. Stanley
Dept. of EECS
University of Central Florida
Orlando, FL 32816-2362 USA
kstanley@eecs.ucf.edu

ABSTRACT

A popular approach to music generation in recent years is to extract rules and statistical relationships by analyzing a large corpus of musical data. The aim of this paper is to present an alternative to such data-intensive techniques. The main idea, called *functional scaffolding for musical composition* (FSMC), exploits a simple yet powerful property of multipart compositions: The pattern of notes and rhythms in different instrumental parts of the same song are *functionally related*. That is, in principle, one part can be expressed as a function of another. The utility of this insight is validated by an application that assists the user in exploring the space of possible accompaniments to preexisting parts through a process called *interactive evolutionary computation*. In effect, without the need for musical expertise, the user explores transforming functions that yield plausible accompaniments derived from preexisting parts. In fact, a survey of listeners shows that participants cannot distinguish songs with computer-generated parts from those that are entirely human composed. Thus this one simple mathematical relationship yields surprisingly convincing results even without any real musical knowledge programmed into the system. With future refinement, FSMC might lead to practical aids for novices aiming to fulfill incomplete visions.

1. INTRODUCTION

An interesting aspect of creativity is that people are often better at discerning and evaluating novelty than at creating it [1, 2, 3]. While researchers define creativity in many ways and assert that creativity *should* be computable, no consensus yet exists on how algorithms should compose creative works [4, 5, 6, 7]. A key challenge for computational models of musical creativity is thus to identify guiding principles that facilitate its creation. The insight in this paper is that it is possible to exploit general mathematical properties of musical relationships and to leverage the intuitive human ability to evaluate novelty in the service of generating accompaniment without musical expertise.

Interestingly, some composers intentionally and explicitly incorporate mathematical transformations into their works. By translating, inverting, and reflecting musical lines they create cannons, fugues, suites, and other styles [8]. However, transformations relating one part of a piece to another are often implicit. For example, Hellegouarch suggests that musicians implicitly incorporated logarithmic and modular transformations long before such operations were formally defined [9]. Similarly, Harkleroad [10] describes the relationship of change ringing to group theory although the musical technique predates the mathematical by centuries. The implicit nature of such transformations thus suggests that the mathematical relationship between musical parts is often intuitive.

An intriguing implication of this *intuitive* aspect of musical transformation is that it may be possible to formulate a *simple* theory that captures some important musical relationships. While much progress in recent years has focused on elucidating abstract relationships [11, 12] or extracting rules from statistical analyses [13, 14], the aim in this paper is to highlight a fundamental and simple insight that requires little data or analysis, in either the construction of the system or its datasets.

The main idea is based on the fact that music is in effect a function of time. That is, the pattern of pitches and the pattern of durations and rests can be expressed together as a vector function of time $\mathbf{f}(t)$ that outputs both pitch and rhythm information. Thus the part played by each instrument in an ensemble piece is the output of such a function. By casting instrumental parts as functions, the problem of accompaniment is illuminated in a useful light: Given an existing part $\mathbf{f}(t)$, the problem of formulating an appealing accompaniment translates to the problem of searching for accompaniment $\mathbf{g}(t)$ such that $\mathbf{g}(t)$ complements $\mathbf{f}(t)$.

Computationally, this problem is difficult if $\mathbf{g}(t)$ is sought *independently* of $\mathbf{f}(t)$ because the space of possible such functions is infinite and unstructured. Thus one approach is to enumerate a comprehensive set of *rules* that in effect describe how $\mathbf{g}(t)$ is constructed [15, 16]. However, the insight in this paper is simpler and interesting because it requires no musical analysis: The function $\mathbf{g}(t)$ can be constrained to a promising set simply by searching instead for $\mathbf{h}(\mathbf{f}(t))$, where $\mathbf{h}$ outputs the accompaniment instead of $\mathbf{g}(t)$. In other words, we can exploit the fact that there must be a *functional relationship* between the accompaniment and preexisting parts, which are thus exploited as a kind of *scaffold*. For example, the scaffold might contain bass and vocal parts, and the functionally-related accompaniment could be a guitar part. By simply searching for a function of the scaffold instead of a raw function of time, the accompaniment becomes constrained by the structure and contours of what is already written. The result is a pattern in time that in effect inherits the human style and character already in the scaffold, without any further analysis or rules. This approach, introduced in this paper, is called *functional scaffolding for musical composition* (FSMC).

A first step toward demonstrating the potential of FSMC is implemented in an interface that allows a human user to direct a search through the space of functional transformations of preexisting scaffolds. Each such transformation generates an accompaniment that the user rates. These ratings then drive a process called *interactive evolutionary computation* (IEC) that in effect allows the user to *breed* new accompaniments. Because each candidate in the search exploits the functional relationship between melody and harmony, the search quickly yields plausible accompaniments that inherit the human essence of the scaffold. In fact, participants in a listener study could not determine whether accompaniment generated through this method is computer-generated or not, even though no other musical knowledge is provided to the system, suggesting the significance of the insight behind FSMC. This result complements that in a companion paper [17], which focuses instead on establishing that the progression of accompaniments evolved through IEC indeed improves in quality.

A further interesting result of the research in this paper is that the functions that express the most convincing accompaniments are often surprisingly simple, implying that the veneer of complexity in the interplay between different musical parts may often be misleading. Perhaps in some cases our appreciation of rich musical tapestry is in its hidden simplicity, which FSMC can expose explicitly. On a practical level, the relative success of such a simple insight represents a possible first step towards more effective computational assistance in musical composition.

2. BACKGROUND

This section places FSMC in the context of other approaches to generating music and reviews a predecessor to FSMC called NEAT Drummer, which established the general principle of functional scaffolding.

2.1 Approaches to Music Generation

An important aspect of many traditional approaches to music generation is that they exploit musical corpora to discover relationships in the data to guide decision-making. For example, Ponsford et al. [18] learn probabilistic n-grams through analyzing local harmonies in a corpus. MySong also uses a hidden Markov model to generate accompaniments [19]. Chuan and Chew [20] combine statistical extraction from a corpus with known musical rules to produce songs in a particular style. While these models yield notable results, they require a significant database that must be carefully constructed. The idea in this paper is to show that accompaniment can be produced simply, without any prior data or analysis.

In particular, the approach in this paper is based on *interactive evolutionary computation* (IEC), whereby a human collaborates with the computer to explore a space of candidates [21]. An example of IEC in music generation is GenJam, a jazz improvisation tool, which learns melodic measures and phrases through IEC [22]. By incorporating human evaluators, IEC helps reduce the need for building and analyzing data from a corpus. By further adding the idea of scaffolding in this paper, IEC is constrained to promising candidates.

2.2 NEAT Drummer

FSMC builds on previous work by Hoover et al. [23] and Hoover and Stanley [24] on NEAT Drummer, a system that creates percussion patterns for existing compositions. The drum generator "hears" the original parts in a MIDI composition simultaneously, and transforms these parts into a drum pattern that follows the contours of the original song. This transformation occurs through a function represented as a compositional pattern producing network (CPPN), which is a special type of artificial neural network (ANN) [25] that can take an arbitrary topology and wherein each neuron is assigned one of *several* activation functions.

NEAT Drummer users can explore drum patterns through IEC with NeuroEvolution of Augmenting Topologies (NEAT), a method for growing and mutating CPPNs [26]. Unlike traditional ANN learning, NEAT is a policy search method, i.e. it *explores* accompaniment possibilities rather than optimizing toward a specific target. While NEAT Drummer showed that the idea of functional scaffolding (implemented through CPPNs) can produce credible percussion accompaniment, it left open the question of whether such an approach can produce complete harmonization, which is the aim of this paper.

3. APPROACH: FUNCTIONAL SCAFFOLDING FOR MUSICAL COMPOSITION

Extending the idea in NEAT Drummer beyond just percussion, FSMC generates complete harmonies from existing compositions. These compositions form the scaffold from which accompaniments are built. However, unlike in NEAT Drummer, these scaffolds include rhythmic information *and* pitch information, thereby providing the foundation for harmonization.

To understand the idea behind FSMC consider that if different instrumental parts in the same composition were not related to each other at all, they would sound inappropriate together. Thus there is some relationship between different parts in the same piece. In effect, this *relationship* can be conceived as a *function* that describes how one part might be transformed into another. That is, theoretically there exists a function that can transform one sequence of notes and rhythmic information into another. If that function is simple then the relationship between the parts is more easily discernible than if the function is complex. Yet in any case, the important point is that there is *some* function that relates these parts to each other. The idea in FSMC is to exploit this fact by literally *evolving the function* that relates one part to another. That way, instead of searching for a sequence of notes, FSMC can search for a transforming function that bootstraps off the existing parts (i.e. called the scaffold) to generate the accompaniment. In effect, FSMC is the hidden function that relates different parts of a composition to each other.

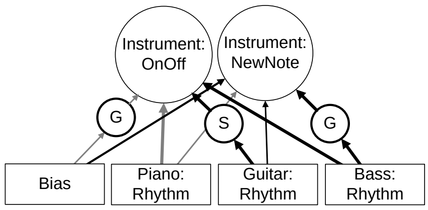
FSMC thus represents accompaniment as a *function* that transforms pitches and rhythms from input tracks (called the scaffold) into a temporal pattern interpreted as the accompaniment. In particular, this transforming function is encoded in FSMC by CPPNs [25], as explained in the next section. Outputs from CPPNs are interpreted as accompaniments that thereby follow contours of the original song. Users then interactively explore the space of such functions for personalized accompaniments through IEC.

3.1 Functional Relationship Representation

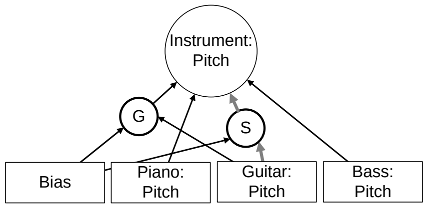
FSMC divides each musical part into a pitch pattern and a rhythm pattern, both of which are represented by separate CPPNs (figure 1). While CPPNs themselves are not essential to FSMC, they serve as convenient representations for exploiting the functional relationships between parts of a piece. The particular idea of separating pitch and rhythm follows a tradition in other approaches to music generation [6, 27]. The rhythm network, which extends the CPPN representation in NEAT Drummer, is shown in figure 1a. It has a set of scaffold inputs from the original composition (i.e. before accompaniment is added) and two output nodes for each instrument in the accompaniment: *OnOff* and *NewNote*. OnOff decides volume and whether or not the note will play. If the OnOff output returns a value below a given threshold, the accompaniment line rests at that tick. If OnOff indicates that a note is to be played, NewNote decides whether the note will be re-struck or sustained. In partnership with the rhythm CPPN, the pitch CPPN in figure 1b sees the pitches of instruments in the scaffold and decides accompaniment pitch with a single output. Viable pitches are discretized into bins that correspond to the given key and the network thereby plays the pitch closest to its output. The CPPNs in figure 1 act just like ANNs with weighted connections and hidden neurons that transform the scaffold input at the current timestep into rhythm and pitch accompaniment.

The CPPN representation in figure 1 thus in effect implements the idea of functional scaffolding. The CPPN is itself just a formalism for specifying a function that can be artificially evolved. The inputs to the CPPNs are the pattern of notes and durations within the scaffold and the outputs are the accompaniment. In this way, the CPPN is literally a function of the scaffold that transforms it into a functionally-related accompaniment pattern.

Figure 2a shows an example of a temporal pattern that is input into the rhythm CPPN. This scaffold is four measures of a repeating quarter-quarter-half note motif. To impart a sense of time within a note, when a note begins, an attack spike is sent to the network for that particular instant in time. This spike decays linearly over time for the duration of the scaffold note. This spike-decay representation of time ensures that the position *within* the particular note is known to the rhythm network at any given time, thereby providing rhythmic context from the scaffold to the accompaniment. Thus it can output patterns based on the rhythmic information in the scaffold. Simultaneously, the pitch from the scaffold at each discrete instant in time is sent modulo 12 to the pitch CPPN (figure 2b), whose output is converted to one of eight pitches in the specified key.



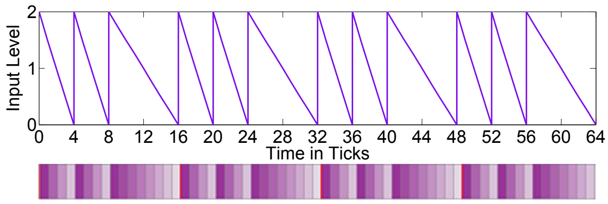
(a) Rhythm CPPN



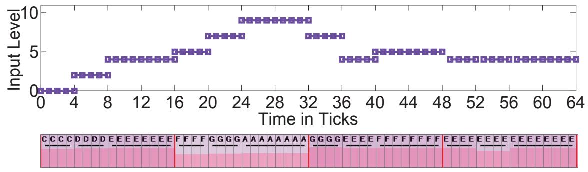
(b) Pitch CPPN

Figure 1. How CPPNs Transform the Input Scaffold. The rhythm CPPN in (a) and pitch CPPN in (b) together form the accompaniments of FSMC. The inputs to the CPPN are the scaffold rhythms and pitches for the respective networks and the outputs indicate the accompaniment rhythms and pitches. Each rhythm network has two outputs: *OnOff* and *NewNote*. The OnOff node controls volume and whether or not a note is played. The NewNote node indicates whether a note is played or sustained at the current tick. If OnOff indicates a rest, the NewNote node is ignored. The pitch CPPN output decides what pitch the accompaniment should play at that particular instant of time. The internal topologies of these networks, which encode the functions they perform, change over evolution.

The hidden nodes in the CPPNs depicted in figure 1 are added by mutations that occur over the evolutionary process. They in effect increase the complexity of the transforming function by adding intervening nonlinearities. For example, the Gaussian function (depicted as a “G”) introduces symmetry (i.e. such as the same sequence of notes ascending and then descending) and the sigmoid (depicted as “S”) is nonlinear yet asymmetric. As in a neural network, the connections are weights (i.e. coefficients) that are multiplied by their inputs. By accumulating such transformations, the relationship between scaffold and accompaniment can become more complex. In effect, the CPPN and its inputs provide the *functional scaffolding* in FSMC. The next section explains how such preferences are conveyed through the evolutionary process.



(a) Spike-Decay Rhythm Representation



(b) Pitch Representation

Figure 2. **CPPN Input Representation.** The spike-decay representation for rhythms is shown in (a) and the pitch representation is shown in (b). Both such inputs are depicted in two ways: The first is a continuous-time graph that shows decaying spikes for rhythm and the pitch level for pitch. The second is a discrete-time representation of what is actually input into the network at each discrete timestep, which is represented by darkness for rhythm and height for pitch. Because the network samples time discretely, results in this paper are also depicted in the discrete-time format. In this way, this figure gives a sense of exactly what the CPPN "hears" (for each instrument in the scaffold) as it generates accompaniment.

3.2 Exploring Functional Space

Through IEC, the FSMC accompaniment gains its musical character from a combination of the human breeder and the information in the scaffold, rather than through a corpus. In the iterative IEC process, the FSMC interface presents a population of candidate accompaniments to the user who then rates each candidate. Pieces are judged good, mediocre, or bad. The next generation then presents to the user accompaniment candidates from CPPNs derived from those judged best in the prior generation. The new generation is created by mutating and mating network weights from the parent CPPNs and occasionally through mutating network structure by adding or subtracting parts of the network, following the NEAT method [26].

Users also influence the accompaniment by choosing CPPN inputs, i.e. to which instruments in the scaffold the CPPNs should listen.

4. EXPERIMENT

The experiment is divided into two parts: First accompaniments are evolved for two songs and then a listener study assesses how convincing the accompaniments are.

4.1 Accompaniments

To demonstrate the capabilities of FSMC, accompaniments were evolved for two pieces in this paper. The program and accompaniment CPPNs will be released in summer of 2011 at `http://eplex.cs.ucf.edu/fsmc`. The two pieces are folk songs (Nancy Whiskey and Bad Girl's Lament) originally arranged in MIDI format by Barry Taylor and redistributed with his permission. The MIDI format is convenient because it is easy to convert directly into the FSMC input format (figure 2).

It is important to note that these pieces are chosen for this experiment because they exemplify entirely human compositions that meet a minimum standard of recognizable quality. While in the future FSMC can potentially generate accompaniment for incomplete compositions by amateur musicians, by starting with pieces that are convincing as complete compositions, it is possible to discern whether the generated accompaniments reduce the human plausibility of the work, or whether they complement it successfully, as would be hoped for such an approach.

The interactive evolutionary process for the two example pieces was guided by the authors. No musical knowledge was applied beyond simply choosing which candidates sounded best. The process proceeded as follows: A set of ten *random* CPPNs corresponding to an initial population of FSMC accompaniments was first generated by the program. Among these, those that sounded best were selected by the user. From the selected candidates a new generation of CPPNs was created that are offspring (i.e. mutations and crossovers) of the original generation. This process of listening to candidates, selecting the best, and creating new generations was repeated until a satisfactory accompaniment appeared. While user input is an important aspect of this process, no session lasted more than 12 generations (i.e. no more than 12 preference decisions were ever made), highlighting the overriding importance of the FSMC relationship to constraining accompaniments to a reasonable set of candidates. Thus, interestingly, in contrast to data-intensive approaches, the only knowledge needed to generate accompaniments through this approach is imparted in ten to 15 clicks of IEC.

Accompaniments are evolved with a CPPN mutation rate and crossover rate of 0.3. The NewNote threshold is also 0.3. Furthermore, when the OnOff output in the rhythm network (which also indicates volume) falls below 0.3, no note is played. The next section explains a study designed to assess the results.

4.2 Listener Study

To gain insight into the potential of the approach, the results in this paper are assessed through a listener study in which anonymous participants were asked to rate examples with and without FSMC accompaniments. The key focus in the study is on whether the fact that a computer is involved in generating some of the examples can be discerned by the listeners. Thus the survey is a kind of *musical Turing Test*. This perspective is interesting because FSMC is based on no musical principle or theory other than establishing a functional relationship; if such a minimalist approach can generate plausible accompaniment it suggests that the theory behind it is at least promising.

A total of 66 listeners, all of whom are students in a diversity of majors at the University of Central Florida, participated in the study. The full survey, including the human compositions, is provided at `http://eplex.cs.ucf.edu/fsmc/smc2011/survey`. The aim is to discover whether the accompaniments sound either natural or computer-generated. Participants are asked to rate five different MIDIs by answering the following question:

> Based on your impression, how likely is it that any of the instrumental parts in the musical piece found at the following link, <link>, were composed by a computer? "Composed" means that the computer actually came up with the notes, i.e. both their pitch and duration, on its own. (1 means very unlikely and 10 means very likely).

The participants rated a total of five MIDIs: (1) an obviously computer-generated control (which helps to establish that participants understand the question), (2) a version of Nancy Whiskey with a computer-generated accompaniment, (3) fully human-composed Chief Douglas' Daughter, (4) fully human-composed Kilgary Mountain, and (5) a version of Bad Girl's Lament with a computer-generated accompaniment. Thus the main issue is whether participants judge piece 2 and piece 5, which have accompaniments evolved with FSMC, as distinguishable from piece 3 and piece 4, which are entirely composed by humans.

5. RESULTS

This section begins with an analysis of the two evolved FSMC accompaniments and then presents the user study. All music discussed in this section, both with and without evolved accompaniments, can be heard at `http://eplex.cs.ucf.edu/fsmc/smc2011`.

5.1 Accompaniments

Figure 3 shows results after two generations of evolving accompaniment for Nancy Whiskey and 12 generations of evolving accompaniment for Bad Girl's Lament. The low number of generations necessary to obtain these results is a result of the strong bias provided by FSMC towards generating accompaniments related to the scaffold. A key issue in understanding the results is the functional relationship between scaffold inputs and CPPN outputs over time, which gives a sense of the implication of linking these parts functionally. To help visualize this relationship, the top line of figure 3a and 3b contains a series of rectangles read from left to right that represent the CPPN output at that particular tick. Rectangle shading indicates pitch and volume: *Darker* color shading represents louder notes while *taller* shading indicates higher pitch. Note names are also written in bold at the top of each rectangle. Notes can be sustained from previous ticks, re-struck, or silenced. A thick horizontal line crossing the borders between ticks indicates a sustain while gaps with thin horizontal lines (slightly lower than the thick lines) indicate rests.

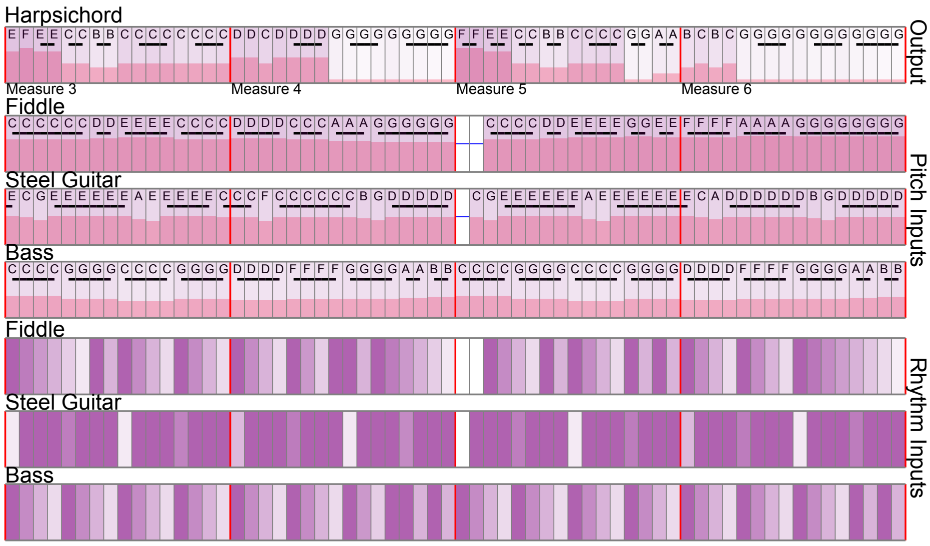
Figure 3 also shows both the rhythm and pitch *inputs* (i.e. the scaffold) to the CPPN. It is important to note that rhythm inputs represent the special spike-decay rhythm format introduced in figure 2a while pitch inputs are simply pitch levels, as in figure 2b.

Figure 3a shows four measures, numbered 3, 4, 5, and 6, of generated accompaniment for the MIDI scaffold, Nancy Whiskey. To also provide perspective in musical notation, figure 4 shows measures five and six of the score. The fiddle, steel guitar, and bass from Nancy Whiskey are input to both the pitch and rhythm networks; the output is a harpsichord accompaniment that inherits pitch and rhythm relationships from the scaffold. One salient such relationship is at the endings in measures four and six where a G is played in the fiddle part. Although the accompaniment does not always follow the fiddle, at this point the output accompaniment also plays G at the same time, although an octave lower. However, while it follows the pitch, the accompaniment varies the rhythm at this part, borrowing rhythmic elements from the other instruments in the scaffold, thereby further differentiating itself from the fiddle. In totality, the accompaniment incorporates pitch and rhythmic elements from all three scaffold instruments while also varying and combining them in novel ways, yielding an original pattern that complements the whole.

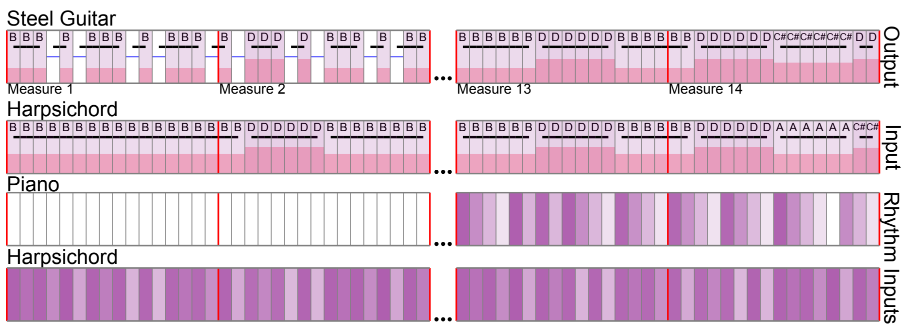
FSMC can be applied to any scaffold. An additional evolved accompaniment for a different scaffold, Bad Girl's Lament, is shown in figure 3b. This image shows measures 1 and 2 of the accompaniment in the first section followed by measures 13 and 14. In measures 1 and 2, the pitches follow the harpsichord input in the pitch network exactly, but the rhythms are different. In the harpsichord part of the scaffold, each note is held until the next pitch sounds. Thus there are no rests or note rearticulations. However, the accompaniment adds new flair to these pitches by inventing a rhythm that is influenced by the rhythm of the harpsichord in the rhythm network. In the generated rhythm, the note is held for three ticks, rests on the fourth tick, plays on the fifth tick, and rests on the sixth tick. This pattern repeats throughout measures 1 and 2. Measures 13 and 14 in figure 3b show how pitches in the accompaniment sometimes also deviate from the scaffold input. Notice that when the harpsichord input plays an A, the accompaniment sounds a C#. Similarly, the accompaniment sounds D when a C# is heard in the scaffold.

Figure 5 shows the internal structure of the evolved CPPNs that produce these accompaniments. In the Nancy Whiskey rhythm CPPN (figure 5a), each input is connected to the two outputs with different weights. In the pitch network, each input is directly connected to the pitch output with the exception of the second steel guitar input, which connects through a hidden node with a Gaussian function. Recall that output nodes compute a sigmoid function of their input.

Interestingly, the CPPNs for generating Bad Girl's Lament accompaniment are even simpler and did not even evolve any hidden nodes (i.e. additional nonlinearities beyond the sigmoid output functions). The simplest CPPN of all, which is the pitch network in figure 5d, has only a single connection. Such relationships encoded by the CPPNs can also be written mathematically. For example, the Nancy Whiskey rhythm CPPN (figure 5a) computes OnOff $= \sigma(-.35n_1 + 1.34n_2 - 1.76n_3 + 1.46n_4)$ and NewNote $= \sigma(1.01n_1 + 1.70n_2 - .37n_3 + .51n_4)$,



(a) Nancy Whiskey Accompaniment



(b) Bad Girl's Lament Accompaniment

Figure 3. **Accompaniments Generated by FSMC.** Accompaniments evolved for Nancy Whiskey and Bad Girl's Lament are shown in (a) and (b). Each accompaniment can be heard at `http://eplex.cs.ucf.edu/fsmc/smc2011`. The accompaniments are produced by CPPNs that relate the scaffold inputs to the generated outputs. The Nancy Whiskey accompaniment shown in (a) is also shown in staff form for measures 5 and 6 in figure 4. The rhythm and pitch relationships between the scaffolds and accompaniments derive from the functional relationship between them.

where $\sigma(x) = \frac{1}{1+e^{-1.1x}}$, and n_i is node number i. In this way, musical relationships really are being encoded as functions. It is important to understand that the simplicity of these relationships resulted from a process of human selection through IEC that ended when the human was satisfied, which means it reflects the human user's implicit preferences. These results show that simple relationships can be appealing and convincing. In this way, this kind of application can tell us something about the nature of the implicit musical relationships that we appreciate.

5.2 Listener Study

The complete results of this study are shown in table 1. On average, the 66 participants judge the intentionally poor example as significantly more likely ($p < 0.001$ according to Student's t-test) to be computer-generated than any other song in the survey. This difference indicates that participants understand the survey.

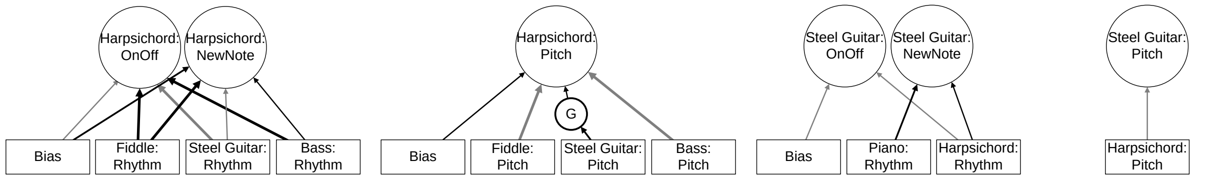
Although the accompanied Nancy Whiskey is judged significantly more likely ($p < 0.05$) to be computerized than the human song Chief Douglas' Daughter, it is not judged significantly more likely than Kilgary Mountain to be computerized. This result indicates that the accompanied Nancy Whiskey can pass the musical Turing test, i.e. people cannot distinguish it from a song that is entirely human-generated.

The Bad Girl's Lament accompaniment is even more difficult for participants to differentiate. It is not judged sig-

Survey Results

MIDI Name	Mean	Std. Dev.
Control	7.82	2.15
Nancy Whiskey with Accomp.	5.45	2.65
Chief Douglas' Daughter	4.32	2.61
Kilgary Mountain	4.86	2.39
Bad Girl's Lament with Accomp.	4.82	2.44

Table 1. **Survey Results (lower means more human-like).** The average ratings and standard deviations for the samples show that FSMC accompaniments can sound human. The Bad Girl's Lament MIDI, which is partly computer-generated, ranks less likely to be computer-generated than the fully human-composed song, Kilgary Mountain, although this difference is not significant.



(a) Nancy Whiskey Rhythm (b) Nancy Whiskey Pitch (c) BGL Rhythm (d) BGL Pitch

Figure 5. **CPPNs Corresponding to Accompaniments.** CPPNs for generating accompaniment for Nancy Whiskey (a and b) and Bad Girl's Lament (BGL; c and d) are shown above. Each accompaniment can be heard at `http://eplex.cs.ucf.edu/fsmc/smc2011`. Line thickness is proportional to weight and gray lines are inhibitory. The simplicity of the CPPNs shows that the functional relationship between different musical parts is often plausible even if it is simple.

Figure 4. **Nancy Whiskey Accompaniment Generated by FSMC in Staff Notation.** Measures 5 and 6 from the score and the generated accompaniment illustrate FSMC accompaniment in staff notation. The top line shows the generated harpsichord output while the next three voices show the fiddle, steel guitar, and bass respectively.

nificantly more likely to be computer-assisted than either of the human pieces, i.e. Chief Douglas' Daughter or Kilgary Mountain. In fact, on average, accompanied Bad Girl's Lament scored slightly *less* likely to be computerized than the entirely human song Kilgary Mountain.

These results validate that evolved accompaniments are at least plausible enough to fool human listeners into confusing partly computer-generated compositions with fully human-composed ones, even though FSMC has almost no a priori musical knowledge programmed into it.

6. DISCUSSION AND FUTURE WORK

The primary contribution of this paper is to show how little prior information is necessary to generate plausible accompaniments. While most generation methods require an extensive corpus to analyze or prior music knowledge [6, 18, 20], this study shows that simple functional relationships are central to musical composition and therefore possible to exploit. The listener study confirms that such accompaniments can be indistinguishable from fully-human compositions by average listeners.

Not only is the insight simple that a functional relationship is foundational to the idea of accompaniment, but the evolved relationships between the accompaniments and the scaffold themselves also turned out to be simple. Nancy Whiskey contains a single hidden node in its rhythm and pitch CPPNs, indicating at most two function compositions. Bad Girl's Lament accompaniment is even simpler. Of course, these results are anecdotal and do not imply that complex relationships would not be appealing in some cases, yet they do raise the intriguing hypothesis that many such relationships could be simpler than they sound intuitively. In the future, it will be interesting to discover in which cases evolution leads to more complex relationships between scaffold and accompaniment.

Interestingly, because the evolved patterns are encoded as transformative functions (i.e. CPPNs), in principle such a function can be applied to a *different* scaffold, which is like transferring an accompaniment "personality" to a new composition. In this way, evolved CPPNs may ultimately apply effectively to multiple pieces of similar genre, which will be explored in the future.

While incorporating at least some musical knowledge into FSMC may ultimately be necessary for widespread application, the contribution so far is to provide a core insight around which further structure can later be built. For the moment, the result that listeners cannot detect the difference between songs with computer-generated accompaniment and those without it suggests that the simple idea of exploiting functional relationships provides a promising starting point for a future research direction.

7. CONCLUSION

This paper provides insight into the relationship between musical parts through a new theory called functional scaffolding for musical composition (FSMC). By representing the relationship between existing parts and a new accompaniment functionally, plausible accompaniments are generated. The resulting pieces with accompaniments are even sometimes indistinguishable from fully-human compositions. The main conclusion is that FSMC provides an alternative to data-intensive approaches to music generation and analysis that nevertheless promises a different kind of insight into the nature of musical accompaniment.

Acknowledgements

This work was supported in part by the National Science Foundation under grant no. IIS-1002507. Special thanks to Barry Taylor for granting special permission to utilize his own MIDI productions of folk music in this work. Barry Taylor originally sequenced Nancy Whiskey and Bad Girl's Lament (without accompaniment).

8. REFERENCES

[1] P. J. Silvia, "Discernment and creativity: How well can people identify their most creative ideas?" *Psychology of Aesthetics, Creativity, and the Arts*, vol. 2, no. 3, pp. 139–146, 2008.

[2] R. J. Sternberg, "The nature of creativity," In *The Essential Sternberg Essays on Intelligence, Psychology, and Education*. Springer, LLC, 2009.

[3] M. L. Maher, "Evaluating creativity in humans, computers, and collectively intelligent systems," In *Proc. of the 1st DESIRE Network Conference on Creativity and Innovation in Design*, 2010, pp. 22–28.

[4] P. N. Johnson-Laird, "Freedom and constrain in creativity," In *The Nature of Creativity: Contemporary Psychological Perspectives*. Press Syndicate of the University of Cambridge, 1988.

[5] R. C. Shank, "Creativity as a mechanical process," In *The Essential Sternberg Essays on Intelligence, Psychology, and Education*. Press Syndicate of the University of Cambridge, 1988.

[6] D. Cope, *Computer Models of Creativity*. Cambridge, MA: MIT Press, 2005.

[7] A. Dorin and K. B. Korb, "Improbable Creativity," In *Proc. of the Dagstuhl International Seminar on Computational Creativity*, 2009.

[8] W. Hodges and R. J. Wilson, "Musical patterns," In *Mathematic and Music: A Diderot Mathematical Forum*. Springer-Verlag, 2002, ch. 5.

[9] J.-C. Risset, "Computing musical sound," In *Mathematic and Music: A Diderot Mathematical Forum*. Springer-Verlag, 2002, ch. 13.

[10] L. Harkleroad, *The Math Behind the Music*, ser. Outlooks. Cambridge University Press, 2006.

[11] S. Cannon, "Maximally smooth diatonic trichord cycles," In *Proc. of the 2nd International Conference on Mathematics and Computation in Music*, 2009.

[12] D. Tymoczko, "Three conceptions of musical distance," In *Proc. of the 2nd International Conference on Mathematics and Computation in Music*, 2009.

[13] C. Rhodes, D. Lewis, and D. Müllensiefen, "Bayesian model selection for harmonic labelling," In *Proc. of the 1st International Conference on Mathematics and Computation in Music*, 2007.

[14] K. M. Kitani and H. Koike, "Improvgenerator: Online grammatical induction for on-the-fly improvisation accompaniment," In *Proc. of the 2010 Conf. on New Interfaces for Musical Expression (NIME 2010)*, 2010.

[15] M. Dahia, H. Santana, E. Trajano, G. L. Ramalho, C. Sandroni, and G. Cabral, "Using patterns to generate rhythmic accompaniment for guitar," In *Proc.of the International Conference on Sound and Music Computing*, 2004, pp. 111–115.

[16] W. F. Walker, "A computer participant in musical improvisation," In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1997, pp. 123–130.

[17] A. K. Hoover, P. A. Szerlip, and K. O. Stanley, "Interactively evolving harmonies through functional scaffolding," In *Proc. of the Genetic and Evolutionary Computation Conference (GECCO 2011)*. New York, NY: ACM, 2011, to appear.

[18] D. Ponsford, G. Wiggins, and C. Mellish, "Statistical learning of harmonic movement," *Journal of New Music Research*, vol. 28, no. 2, pp. 150–177, 1999.

[19] I. Simon, D. Morris, and S. Basu, "MySong: automatic accompaniment generation for vocal melodies," In *Proc. of the Twenty-Sixth Annual SIGCHI Conf. on Human Factors in Computing Systems*. ACM, 2008, pp. 725–734.

[20] C.-H. Chuan and E. Chew, "A hybrid system for automatic generation of style-specific accompaniment," In *4th International Joint Workshop on Computational Creativity*, 2007.

[21] H. Takagi, "Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation," *Proc. of the IEEE*, vol. 89, no. 9, pp. 1275–1296, Sep 2001.

[22] J. A. Biles, "Interactive genjam: Integrating real-time performance with a genetic algorithm," In *Proc. of the 1998 International Computer Music Conference*, 1998.

[23] A. K. Hoover, M. P. Rosario, and K. O. Stanley, "Scaffolding for interactively evolving novel drum tracks for existing songs," In *Proc. of the Sixth European Workshop on Evolutionary and Biologically Inspired Music, Sound, Art and Design (EvoMUSART 2008)*, 2008.

[24] A. K. Hoover and K. O. Stanley, "Exploiting functional relationships in musical composition," *Connection Science Special Issue on Music, Brain, & Cognition*, vol. 21, no. 2, pp. 227–251, June 2009.

[25] K. O. Stanley, "Compositional pattern producing networks: A novel abstraction of development," *Genetic Programming and Evolvable Machines Special Issue on Developmental Systems*, vol. 8, no. 2, pp. 131–162, 2007.

[26] K. O. Stanley and R. Miikkulainen, "Evolving neural networks through augmenting topologies," *Evolutionary Computation*, vol. 10, pp. 99–127, 2002.

[27] P. Johnson-Laird, "Jazz improvisation: A theory at the computational level," In *Representing Musical Structure*, ser. Cognitive Science Series. London, UK: Academic Press Limited, 1991, ch. 9.