A Combinational Creativity Approach to Composing Traditional Irish Reels

Nan Zheng[1] , Bryan Duggan[1]

[1] School of Computing, Dublin Institute of Technology, Kevin St., Dublin 8, Ireland
nan.zheng1@student.dit.ie, bryan.duggan@dit.ie

Abstract. In this paper we describe a system that uses a corpus of 864 traditional Irish reels as input into an algorithm that composes new tunes. The system performs a structural analysis of the tunes in the corpus and also counts n-gram note sequences in the tunes. It then recombines n-gram note sequences together in structures from the corpus to generate new tunes. We further present our evaluation of the generated tunes as performed by 29 domain experts.

Keywords: Irish traditional music, reel, algorithmic composition, stochastic sampling, n-grams.

1 Introduction

The most common forms of traditional dance music are *reels*, *double jigs* and *hornpipes*. Other tune types include *marches*, *set dances*, *polkas*, *mazurkas*, *slip jigs*, *single jigs and reels*, *flings*, *highlands*, *scottisches*, *barn dances*, *strathspeys* and *waltzes* [1]. These forms differ in time signature, tempo and structure. A reel is generally played at a lively tempo and is in 4/4 time (although played and transcribed as 8 quavers in a bar). Most tunes have a structure consisting of two "parts" (of length eight bars each) called the A part and B part. In a double reel, each part is played twice and then the entire structure is repeated up to three times. Typically within each part there is a certain amount of repetition, which usually occurs in half bar phrases.

[2] describes combinational creativity as "*novel combinations of old ideas*". It is clear from any analysis of a corpus of traditional music that there are many examples of combinational creativity. The tunes "The Bag of Spuds" and "Down the Broom" have very similar B parts for example while the tunes "Sleepy Maggie" and "Jenny's Chickens" share very similar A and B parts. In this paper we describe our work in developing a combinational creativity algorithm that can compose new reels by first analysing the structure of tunes from a corpus and then recombining n-grams of notes from the corpus to create new reels. Our system takes advantage of the fact that half bars in tunes from the corpus often occur several times in a tune to create structures that should sound correct. The fact that note sequences in the generated tunes come from n-grams of notes in the corpus should also mean that the generated melodies are plausible. Our system generates any number of new tunes. To test the system, we generated one hundred new tunes using our algorithm and selected nine at

random for evaluation by domain experts. We included one human composed tune for comparison and asked the experts to try and guess which tune was composed by a human.

Section 2 presents background information on the corpus that we used as input to our system and presents the ethnographic context of the corpus. Section 3 of this paper describes related work in the field of algorithmic composition focusing of systems that purport to create folk melodies. Section 4 of this paper describes our algorithm in detail. Section 5 of the paper presents an evaluation of the generated tunes performed by 29 domain experts, while section 6 presents conclusions and future work.

2 Background

Current estimates suggest there are at least seven thousand traditional tunes in existence [3]. One of the main reasons proposed for the existence of such a wide repertoire is the geographic isolation of rural Irish communities in the centuries preceding the twentieth century [4]. It is proposed that many isolated rural communities developed their own repertoire of tunes and that widespread knowledge of a common repertoire did not occur until the publication of catalogues of traditional tunes such as O'Neill's The Music of Ireland in 1903. In his seminal work, O'Neill collected 1850 tunes played by emigrant Irish musicians in Chicago at the time. Several of the tunes in the catalogue are variations of the same tune.

In 1991, the ABC music notation language was introduced by Chris Walshaw [5]. The format was designed primarily for folk and traditional tunes of Western European origin which can be written on one stave in standard classical notation [5]. ABC files are ASCII text files and so can be edited by any text editor, without the necessity for special software. Each file (known as a tune book) can contain multiple tunes. File sizes are typically measured in kilo-bytes and this facilitates easy transmission by electronic means.

Figure 1 is the tune "Contentment is Wealth" in the ABC format. Each tune consists of a header section and a tune body. The header section contains amongst other fields, the title, com-poser, source, tempo, key signature, geographical origin and transcriber [6]. As tunes can have several titles, the title field can be repeated for a given tune.

```
X:11
T:Contentment is Wealth
R:jig
M:6/8
K:Edor
GFG Eed|BAB EFG|FAF DdB|AFD D2f|gfe edB|BAB ~d3|BdB
DFA|GED E3:|
|:ede Beg|bge gfe|dcd Adf|afd fed|ede Beg|bge gfe|BdB
DFA|GED E3:|
```

Figure 1: The tune "Contentment is Wealth" in the ABC format.

The tune body contains the notation for the tune. The body encoding supports such features as ornaments, bar divisions, sharps, flats, naturals, repeated sections, key changes, guitar chords, lyrics and variations. Between 1997 and 2000, a group of musicians under the leadership of Dan Beimborn and John Chambers, undertook a grass roots project to transcribe three of O'Neill's books to electronic format using the ABC music notation language. As copyright had expired on O'Neill's original books, they made their work freely available on the internet [7].

Many of the tunes from O'Neill's books are played differently by musicians today, as is normal with a living tradition. Around the same period (the late 1990's) Henrik Norbeck collected nearly 2000 tunes in ABC format from various sessions and recordings. Again this collection was made freely available on the internet. This collection contains many modern settings of tunes from O'Neill's books [8] Our system uses a corpus of 864 reels in ABC format drawn from Henrik Norbeck's transcriptions.

3 Related Work

[9] describes the CONCERT system. CONCERT is first trained on melodies represented as a sequence of note pitch names, durations and chords. Various corpora were used to train CONCERT including sets of J. S. Bach pieces and traditional European folk melodies. Internally, CONCERT uses a recurrent network architecture that learns to behave as an autopredictor. A melody is presented to it, one note at a time, and its task at each point in time is to predict the next note in the melody.

The author reports that while the system performed well on simple, structured, artificial sequences, the architecture failed to capture global musical structure. He reports that few listeners were fooled into believing that the pieces had been composed by a human and describes the output of his model as "music only its mother could love".

In [10], the authors describe a novel system that uses a neural network trained on one thousand traditional Irish tunes (jigs, reels and slow airs) and that uses Irish rainfall data as input to generate new melodies. They first convert the training set to MIDI and truncate each of the tunes to be less than 128 note events. They used two back propagation neural networks, one for pitch and one for duration and trained the networks to recognise each of the 1000 tunes from the training set. They then forced the normalised rainfall data upon the inputs to the neural networks, which resulted in the networks producing output vectors of 128 note events. As they had data for one year, they generated 365 note sequences. They developed a windowing system that extracted sequences of the generated melodies in order and re-sequenced them to form a playable melody. The authors provide no validation of the quality of the generated melody, but the generated melody was played by the Irish Chamber Orchestra and was chosen to be one of the centrepieces in the Irish Pavilion at EXPO2000 in Hanover. To our ear, the melody sounds atonal and lacking in musical structure however and quite unlike the tunes from the corpus.

In [11] the authors describe a system that identifies long timescale musical structures in MIDI data generated from ABC files. Further, their system uses knowledge learned about musical structure to generate new melodies. Firstly, they take a corpus of 435 reels transposed into the same key and generate MIDI from these tunes. They then use a meter extraction algorithm which returns a series of timelags corresponding to multiple levels in the metrical hierarchy. The melodies and the meters are fed into a special type of recurrent neural network called a Long Short Term Memory Network (LSTM) which is trained to predict over all possible notes at time t using as input the note (and chord) values at time $t - 1$. The authors claim the advantage of using a LSTM network is that it can learn long timescale (global) musical structure. They generate new tunes by presenting the network with the first few notes of a new tune (not from the corpus) and using it to predict the subsequent notes. They subjectively claim that their system generates “new and interesting” melodies, but present no experimental validation of the generated melodies.

Our system uses a simpler approach to generating new melodies that takes advantage of the fact that the input corpus is a text based markup language for musical scores. This facilitates the use of string comparisons in the analysis of the structure of the scores in our corpus. We also present an evaluation by 29 domain experts of the melodies generated using this approach.

4 A Pattern-based Sampling Approach

Figure 2 is a high level diagram describing the processes in our system.

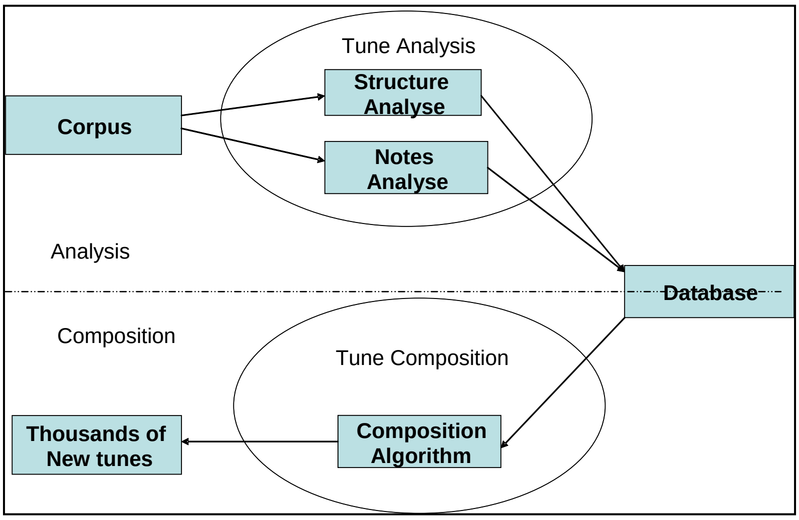

Figure 2: High level diagram of the composition system

In the analysis of tunes in the corpus, the system looks at both the structure of each tune and also the note sequences in each tune. We first apply an algorithm to reveal repeated patterns and these pattern structures will be used later during the stochastic sampling process used to generate new tunes. This is one way to overcome the limitations with purely n-gram models of music composition [12].

For each tune, we create an abstraction of the tune based on reoccurring tokens in the tune. There are various ways in the ABC language to represent repeats and these are considered by the system in generating a structure. Each half bar (4 quaver sequence) is considered to be a token.

```
  BG~G2 BGcG|BG~G2 Bdgd|BG~G2 BdcB|1      | 1 2 |1 3 |1 4
ADFG ABcA:|2 AGFG ABcA||                  | |5 6 |
  ~g3d BGBd|~g2eg faaf|g2gd BddB|ADFG     | 1 2 |1 3 |1 4
ABcA|                                     | |7 6 |
  ~g3d BGBd|~g2eg fa~a2|bgaf gedB|AGFG    | 8 9 |10 11 |12
ABcA||                                    | 13 |5 6 |
  Bdgd Bdgd|Bdgd BG~G2|Ac=fc Acfc|Ac=fc   | 8 9 |10 14 |15
BG~G2|                                    | 16 |7 6 |
Bdgd Bdgd|Bdef ~g3a|bgaf gedB|AGFG        | 3 3 |3 1 |17 18
ABcA||                                    | |17 1 |
```

Figure 3: Original ABC notation of the tune "The Flogging Reel" and the structure of the tune based on reoccurring 4-gram tokens

We assign a number to each unique token in the tune body. The first token number is 1. If the second token differs from the previous token then the number 2 is assigned to it. The algorithm then compares the third token with the previous two and if it is the same token as a previous token, it is assigned the same number. If not, a new number will be assigned and so the algorithm continues until the end of the tune is reached.

Symbol	Interpretation
`BG~G2`	A roll. ~ ignored.
`Ac=fc`	An accidental. = is ignored
`(3BAG dB cAFA`	A triplet, three notes in the time of two. (3BAG is considered to be a block that takes up two places. (3BA should not be separated
`BG~G2`	G2 is a G played with a length of two notes, so that it should not be separated. This means the G2 should be used as one block and takes up 2 places.
`C'ABA`	c' is considered to be one note and should not be separated from the note before it.
`D,ACA`	D, is considered to be one note and not separated from the note before it

Figure 4: ABC Language features that complicate the generation of n-grams

The whole tune is then converted to a sequence of token numbers. This abstracts the tune structure from the notes played in the tune. As some of the structures in the corpus are shared by several tunes we generate roughly 800 tune structures. Figure 3 is an example of the original ABC notation for the tune “The Flogging Reel” and the structure generated by our analysis. The algorithm then stores the tunebook location and name, the number of the tune in the tunebook, the name of the tune and the abstract structure represented by a sequence of token numbers in a database.

We then analyse note sequences in the corpus using n-grams. With different n values, n-gram analysis will generate note combinations for every possible combination of n notes in a tune. There are additional symbols in the ABC language that complicate the generation of n-grams. These symbols are summarised in Figure 4.

Along with the n-gram, the key of the source tune that the n-gram occurred in is also stored. This is because the key of the source tune will determine the actual note combinations in the tune. The key is then an identifier to identify musical phrases that occur in a particular key.

The system is flexible and supports any value of n but currently for ease of composition, the system counts 4-gram musical phrases. With n value equals to 4, there are 220,000 4-grams in the corpus. Figure 5 is an example of n-grams generated from the start of the tune "The Flogging Reel".

ID	GramContent	GramN	Key
181138	BGG2	4	Gmix
181139	GG2B	4	Gmix
181140	G2BG	4	Gmix
181141	BGcG	4	Gmix
181142	GcGB	4	Gmix
181143	cGBG	4	Gmix

Figure 5: Example n-grams generated from the start of the tune "The Flogging Reel"

To generate a new tune, the system first takes a structure at random from the database. Based on the key of the structure, the algorithm fetches an appropriate number of n-grams from the database and fills in an n-gram token into each token in the sequence. For each structure, a user configurable number of new tunes can be generated.

5 Evaluation

In order to evaluate our approach, we generated one hundred tunes and selected nine at random from the generated tunes for testing purposes. To the nine generated tunes, we added one human composed tune. MIDI renderings of the tunes were created so that subjects could listen to the tunes. We created an online survey using phpESP and asked subjects to rate each tune on a scale from one (poor) to five (excellent) in three categories:

1. Originality – To what extent the tunes differed from tunes already in the repertoire of the subject.
2. Aesthetic value – Did the subject like the tune.
3. Correctness – Did the tune sound right or did the tune sound odd.

We asked each subject three additional questions:

1. How long they had been playing music for?
2. Which of the ten tunes was their favourite?

3. Which one of the ten tunes was composed by a human?

We also left a freeform field on the survey and invited subjects to comment on the test. We posted an invitation to complete the survey on two popular web discussion forums used by traditional musicians [13, 14].

Within two days, the survey had been filled out by twenty nine subjects and our invitation had triggered a lively discussion as to which was the human composed tune on one of forums.

The average rating for each of the ten tunes in each of the three categories is presented in Figure 6.

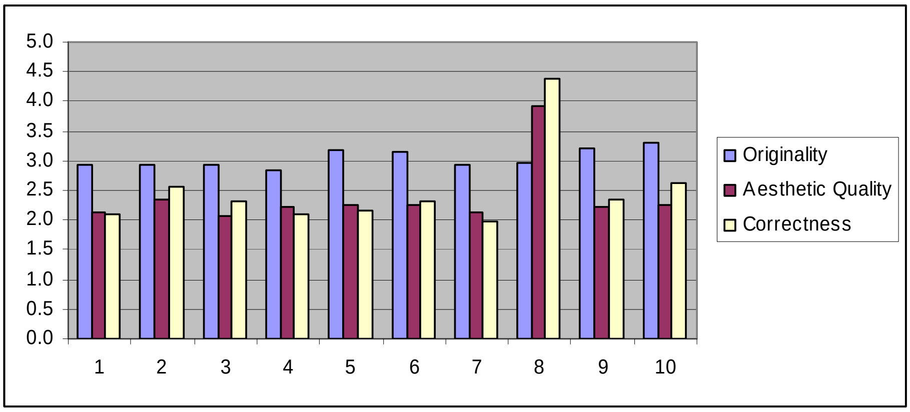

Figure 6: Average results for originality, aesthetic quality and correctness of the ten tunes in the test

On average, subjects rated the computer generated tunes 3 for originality, 2.2 for aesthetic quality and 2.3 for correctness. By contrast, subjects rated the human generated tune on average 3.0 for originality, 3.9 for aesthetic quality and 4.4 for correctness. This was not surprising as many of the computer generated tunes although they repeated in the correct places, contained unusually large pitch intervals that lent the tunes a somewhat atonal quality. This could have been corrected by "improving" the algorithm by filtering large pitch intervals, forcing the melodies to resolve or adding rules, however we judged that the tunes had a unique value due to this atonal quality that might be eliminated if the algorithm was too "constrained".

Most subjects (72%) correctly identified tune eight as being the human composed tune and similarly most subjects (58%) chose the human generated tune as their favourite from the selection. 13% of the subjects judged tune one to be their favourite, which was the favourite tune of the authors.

Most interesting from the survey was the feedback received in freeform comments submitted by subjects. Some of the subjects felt that many of the tunes combined interesting and aesthetically appealing phrases with atonal and unusual phrases. Pitch intervals in some of the tunes were unusually large for traditional tunes, with subjects commenting that many of the tunes had a modern feel reminiscent of modern Scots piping tunes. Many of the subjects commented that had the tunes been worked on by and played by a human rather than the MIDI rendering, the perceived imperfections could have been eliminated.

6 Conclusions

In this paper we presented our system that combined n-grams of notes together from a corpus into tune structures derived from the corpus to compose new tunes. We evaluated the generated tunes and had subjects try and tell the human composed tune from the generated tunes. From our evaluation it was clear that most subjects preferred the human composed tune to the tunes generated by our system. The tunes generated by our system, because they are structurally based on a tune from the corpus, usually repeat in the correct places and so often sound as though there was some planning in their construction. We feel that although many of the tunes have a strange atonal quality we are reluctant to dismiss them, because in our opinion some of the tunes possess interesting chord progressions and are curiously pleasing although they are in many ways different to what most listeners would classify as traditional reels. On the other hand we conclude that our use of random n-grams from the corpus needs further work if we are to improve the perceived correctness and thus the aesthetic quality of the tunes. The tunes used in this test can be listened to here: http://www.bryanduggan.com/phpESP/public/survey.php?name=GeneratedTunes.

References

1. Vallely F. The Companion to Irish Traditional Music: New York University Press; 1999.
2. Boden MA. Dimensions of creativity. Cambridge, Massachusetts: MIT Press; 1996.
3. Wallis G, Wilson S. The Rough Guide to Irish Music. First Edition ed. London: Rough Guides; 2001.
4. Vallely F. Flute Routes to 21st Century Ireland: National University of Ireland; 2004.
5. Walshaw C. The ABC home page. 2007 [cited; Available from: http://www.walshaw.plus.com/abc/
6. Mansfield S. How to Interpret ABC Notation. 2007 [cited; Available from: http://www.lesession.co.uk/abc/abc_notation.htm
7. Chambers J. O' Neills Books. 2007 [cited; Available from: http://trillian.mit.edu/~jc/music/book/oneills/
8. Norbeck H. ABC Tunes. 2007 [cited; Available from: http://www.norbeck.nu/abc/index.html,
9. Mozer MC. Neural network music composition by prediction: Exploring the benefits of psychoacoustic constraints and multiscale processing. Connection Science 1994.
10. Fernström M, Griffith N, Taylor S. BLIAIN LE BAISTEACH - SONIFYING A YEAR WITH RAIN. In: Proceedings of the 2001 International Conference on Auditory Display; 2001 July 29-August 1, 2001; Espoo, Finland; 2001.

11. Eck D, Lapamle J. Learning Musical Structure Directly from Sequences of Music. In: University of Montreal, Department of Computer Science, CP 6128, Succ. Centre-Ville, Montreal, Quebec H3C 3J7 Canada; 2006.
12. Conklin D. Music Generation from Statistical Models. Proceedings of the AISB 2007.
13. Chiff&Fipple. Chiff & Fipple Forums. 2007 [cited; Available from: http://chiffboard.mati.ca/
14. thesession.org. The session.org Forums. 2007 [cited; Available from: http://www.thesession.org