The Checkerboard of Tunes

Andrew Crompton

School of the Arts, University of Liverpool, UK; andrew@cromp.com

Abstract

A physical model was made of all the tunes in Deny Parsons' *Directory of Tunes*. The Directory encodes tunes as a sequence of the letters, U, R, or D, according to whether the notes go up, repeat, or go down, with respect to the previous note. If we substitute +1, 0, -1 for U, R, D respectively, then let the Parsons number, S(n), be their sum up to the nth note. When the number of tunes with particular values of S for n = 1, 2, ...16, were counted for all the tunes in the Directory the results, seen in model form, showed a surprising checkerboard pattern. It is argued that the checkerboard is musically significant and arises because the distribution of repeating notes in tunes is non-random.

What Does a Tune Look Like?

For those of us who are able to read a score, a tune looks the way it is written. Even those of us who cannot read music can grasp the shape of a tune because high notes map to marks higher up the stave. This simple correspondence is the basis of animated graphics like those found on YouTube in which notes, represented by blocks of glowing color, flow right to left, rising and falling with the pitch. The combination of light and music is a kind of synesthesia that gives the illusion that you can read music; rather like film subtitles allow you to pretend you can speak another language. These animations show us that music lies between order and disorder, for as soon as a pattern seems to emerge it dissolves in unexpected ways. Walter Pater wrote that 'All art constantly aspires towards the condition of music' [4], and it was probably this quality of abstract indescribability that he was thinking of. Are there any other ways to show how music is poised between order and chaos? An attempt was made to represent tunes en masse as a physical model in the hope of seeing if they have any statistical regularity that distinguishes them from random sequences of notes.

The model was based on a simple way of identifying tunes invented by Denys Parsons for his *Directory of Tunes,* first published by the mathematician and philosopher G. Spencer Brown in 1975, [5]. It was the subject of an article in *New Scientist*, [6], and second hand copies commanded high prices until it was reprinted in 2002. Unlike older dictionaries based on ordinary music notation that required the reader to transpose the melody into a particular key to identify it, Parsons' Directory can be used by almost anyone. His method, shown in Figure (1), ignores key, interval size, duration and rhythm. Notes are coded U, D, or R, according to whether they go up, down, or repeat, with respect to the previous note. A typical entry looks something like this: *DDUURRDRRUURD, (Mary had a Little Lamb). The initial letter tells you whether the second note goes up, down, or is a repeat with respect to the first note which is represented by the asterisk:

He lists tunes to 13 or 15 places. Most entries differ from their neighbor after between 4 and 11 places, rarely are all 15 letters needed to identify them. Figure 1 shows the score for 'Mary had a Little Lamb' above its Parsons code, and below them a representation of the tune progressing from left to right by making pawn-like moves. This simple tune takes a fairly conventional route; others rise or fall more sharply.

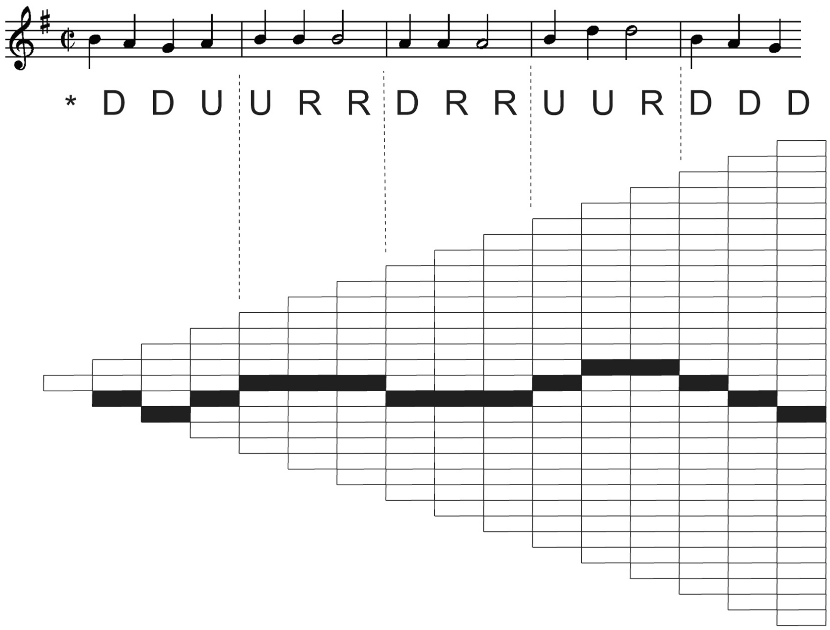

Figure 1: *Code for 'Mary had a Little Lamb', from the Popular Tunes section of the Directory, shown graphically as a series of steps in Parsons space.*

Only one theme in the Directory, by Saint-Saëns, descends for fifteen consecutive notes to reach the cell in the bottom right of the triangular table, and only two tunes ascend for fifteen notes to reach the cell in the top right. Other cells on the last column are reached by multiple routes.

If you were to draw a path like this for every tune in the Directory, which routes would be the most travelled? This question was answered by digitizing the Directory then converting the code for each tune into a series of numbers by substituting +1, 0, -1 for U, R and D respectively. Let their sum up to the nth note, S(n), be called the Parsons' number of that note, with S(1) = 0. For example, after 16 notes 'Mary had a Little Lamb', descends to cell S(16) = -2. The values of S(n) for n = 1, 2, ...16, were counted for each of the 7523 classical themes in the Directory whose code was given to 15 places. The results are shown numerically in Figure 2, in which it can be seen that a total of 553 classical tunes also reached the same cell as 'Mary had a Little Lamb'. Observe that the numbers in the dark cells are always higher than their neighbors. This unexpected checkerboard pattern can be seen in physical form in the stereolithographic model, photographed in Figure 3, in which the height of the cells is proportional to the number of tunes passing through them. Models for pop tunes and individual composers, in Figures 4 and 5, are similar.

The models show a preference for the second note to be up rather than down, an initial asymmetry which diminishes as the tunes progress. Other than this blemish the models all show the same graceful sweep as tunes expand into Parsons' space. No composer has written enough tunes on their own to reach the extremities of Figure 1, but models for Bach, Beethoven and Mozart individually show the checkerboard persists wherever their tunes reach, with the pattern being stronger and smoother for Bach than Mozart, Figure 5. The checkerboard pattern for these composers appears qualitatively similar to that of the whole database. What does the pattern mean?

S(n) \ n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
15																2
14															3	2
13														4	1	4
12													5	2	6	1
11												9	4	12	3	16
10											12	5	14	5	31	6
9										18	6	23	9	49	10	55
8									35	9	40	11	79	25	82	28
7								62	16	69	28	118	39	108	53	127
6							125	21	136	56	167	60	164	79	209	97
5						264	55	290	94	289	107	317	135	341	146	343
4					481	100	493	161	506	171	517	206	547	265	517	269
3				1040	237	867	323	841	320	829	291	769	378	726	377	708
2			1862	453	1449	509	1305	509	1162	473	1108	504	970	485	908	485
1		3825	889	2221	795	1945	696	1554	647	1435	717	1302	621	1149	574	1001
0	7523	1338	3157	980	2148	801	1872	754	1627	741	1376	708	1297	655	1172	625
-1		2360	503	1994	681	1654	654	1586	703	1333	656	1208	632	1157	637	1076
-2			1112	283	1312	417	1237	477	1158	539	1115	560	1018	549	946	553
-3				552	155	758	251	827	314	847	383	814	407	780	462	751
-4					265	76	398	151	551	222	562	258	611	299	611	362
-5						132	43	218	102	362	147	395	180	445	218	475
-6							71	26	116	52	207	82	246	116	305	145
-7								46	17	54	34	122	59	174	74	225
-8									19	9	43	20	74	31	112	52
-9										15	1	28	11	42	22	70
-10											6	0	20	6	28	12
-11												4	0	16	3	21
-12													3	0	11	2
-13														3	0	9
-14															2	0
-15																1

Figure 2: *Parsons number, S(n), for n = 1,2,..15 for 7523 classical tunes and themes all 16 notes long. Starting in the left hand box, tunes progress from left to right by making pawn-like moves. Cells are marked with the number of tunes passing through them. Each column adds up to 7523.*

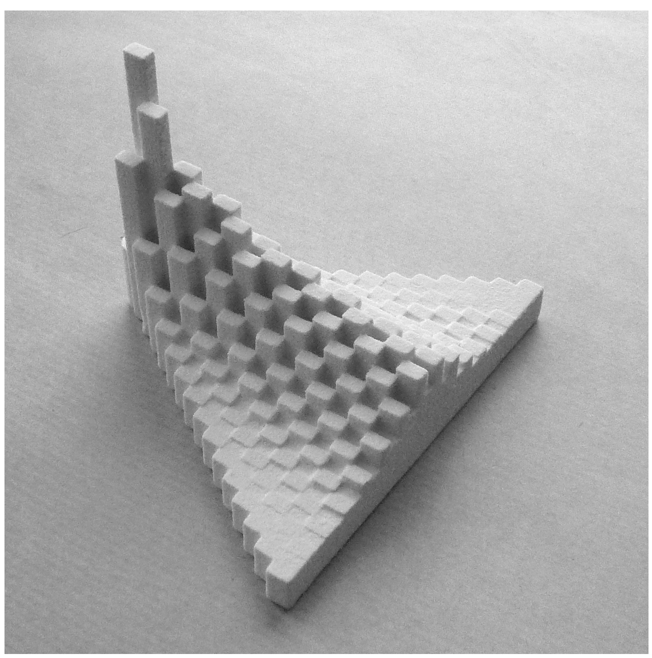

Figure 3: *Stereolithographic model of Figure 2 showing 7523 classical themes. The height of cells is proportional to the number of tunes passing through them.*

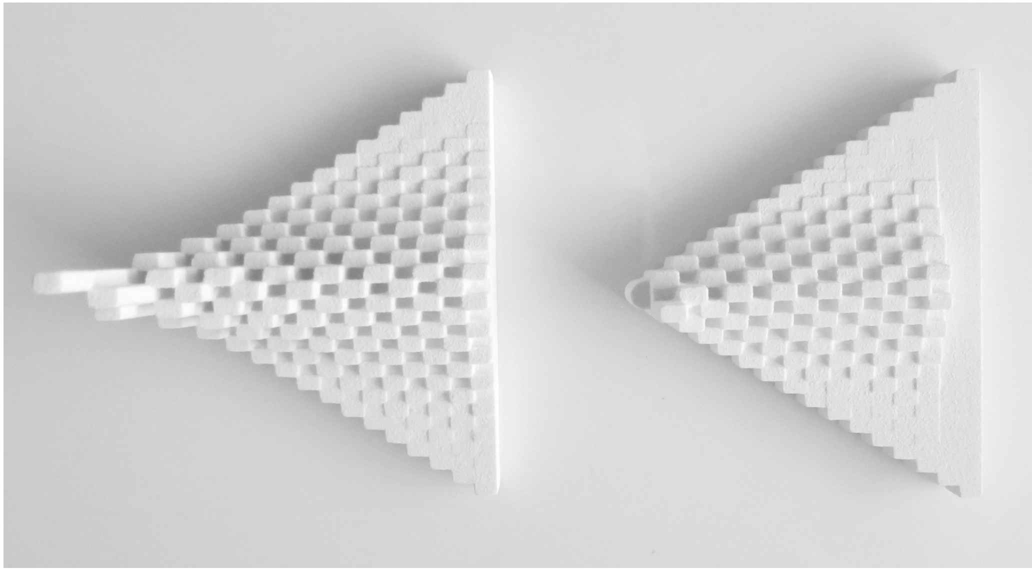

Figure 4: *Left: 7523 Classical tunes and themes 16 notes long. Right: 3661 pop tunes 14 notes long.*

Figure 5: *Left to right: 634 Mozart, 457 Beethoven, and 543 Bach tunes.*

The Origin of the Checkerboard

The origin of the checkerboard pattern lies in the way the data is presented but, as will be shown, it is also musically meaningful. If a tune has no repeating notes then it is confined to the dark cells in Figure 1 like a bishop in chess: tunes with no repeating notes must lie on the checkerboard. There are 2859 classical tunes with no repeating notes in the Directory and they make a perfect checkerboard pattern. When the other 4664 tunes that happen to contain repeating notes are modeled on their own, their checkerboard pattern fades after the first two or three notes. The pattern we see in the models is a combination of two sets of data, a perfect checkerboard standing on a smooth curved surface.

Tunes step up, down, or stay the same, with the following transition probabilities: p(U) = 0.42, p(R) = 0.13, p(D) = 0.45. In an experiment random sequences of U's, D's, and R's were constructed with these values. The checkerboard for these pseudo-tunes faded after the first four or five notes. This tells us that the U's, D's and R's are not distributed randomly in real music. What is this non-random property that tunes possess which causes them to lie on the checkerboard?

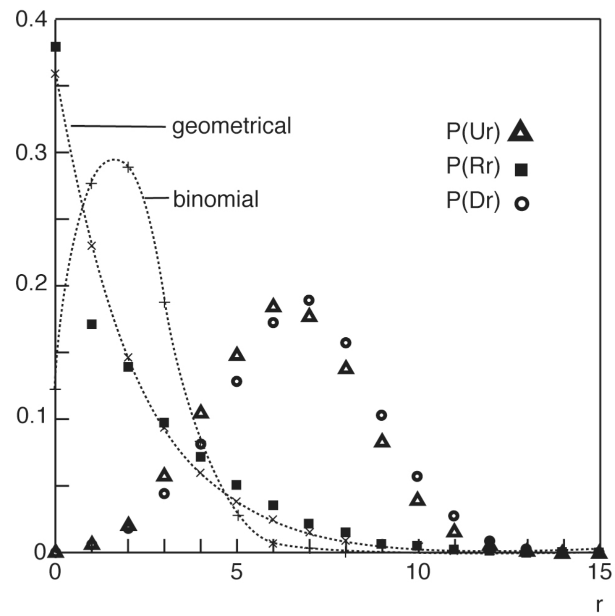

Figure 6: *The probability, P(Ur), P(Rr), P(Dr), that there be r instances of U, D, or R in a code for values of r = 1,2..15. The point in the top left shows that 38% of all tunes have zero repeats. Dashed lines show how values of P(Rr) would lie if distributed geometrically, × points; or binomially, + points.*

The actual distribution of U's, D's, and R's is shown in Figure 6. Although the U's and D's are close to binomial, they cannot be exactly so because their complement, the distribution of R's, is not. The distribution of R's has a long tail with zero repeats being the most likely. It looks approximately geometrical, but there are fewer tunes with a single repeated note than any natural curve passing through the points in Figure 6 would require. The distribution of repeats would be binomial if the music of each composer could be modeled as notes chosen randomly from a set with different relative frequencies, in which case the number of repeats would be lowest when those frequencies were equal thus maximizing the entropy of all notes. Evidently true music is not like this. Note, however, that composers approach this state of randomness to greater or lesser degrees; compared to Mozart, Bach is more uniform in his preferences for particular intervals, [1], and is less likely to repeat notes in a piece of music, [2].

Tunes are close to random, but not perfectly so. An important difference is that repeating notes are not distributed randomly. They occur with a probability of 0.13, if they were distributed binomially with this probability about 12.6% of 16-note sequences would have no repeats in them. In fact 38% of tunes have no repeating notes, almost three times more than chance would suggest. It is this surplus of tunes with no repeats that gives rise to the checkerboard pattern.

The theory that the pattern is caused by the relative frequency of U's, D's, and R's and not on their order was further tested by permuting, randomly and independently, the code for each tune to create a set of codes with the same frequencies of U's, D's and R's as those in Figure 2, though no longer corresponding to tunes. These gave rise to versions of the table in Figure 2 with different values except for the first and last columns. In a hundred trials they were found to have a less pronounced checkerboard structure according to a simple measure of average differences between white and dark cells. These simulations all fell short of the true tune checkerboard. This shows that the order of the U's, D's and R's in the tunes is significant, although why this makes a difference is unknown.

Another way of looking at repeating notes might be to say that they are just parts of long notes arbitrarily divided. This suggests an even simpler way of encoding tunes using only U's and D's, resulting in a binary rather than a base-three code. A test based on Parsons' data showed that the median number of U's, D's and R's needed to identify a tune is nine, and that 5% of the tunes could not be distinguished after 15 notes. Ignoring R's makes matters worse, the median length of code before a difference occurs increases to 12 notes and 20% of tunes are not unique after 15 notes. This shows why Parsons was wise to use U, D and R instead of simply using U and D. It also tells us that repeating notes are one of the ways that we can tell tunes apart.

Summary and Conclusions

The checkerboard results mainly from the unlikely distribution of repeating notes in tunes. It seems that, (i) composers favor tunes with no repeats and there are fewer tunes with a single repeat than one would expect if notes were chosen randomly, (ii) once the decision to use repeats is made, where they occur is significant, and (iii) recognizing repeating notes is one of the ways we can identify tunes. Obviously too many would be boring, although long sequences of repeated notes do exist in music; for instance, *The Boy Next Door*, a song by Martin and Blane sung by Judy Garland in the movie *Meet Me In St Louis,* contains a sequence of 24 repeated notes whose lyric is: *For I live at fifty-one-thirty-five Kensington Avenue, and he lives at fifty-one-thirty-three*. Despite this Parsons does not include RRRRRRRRRRRRRRR in his list of tunes even though it may be the oldest tune of all.

No definition is known of what tunes are, [3], and no recipe is known for their creation. Sadly, algorithmic models rarely manage to create aesthetically satisfying music even though success would be commercially valuable. The statistical property of tunes shown here could be used to assess any such algorithm by seeing if it produced tunes with a checkerboard pattern, but this would be after the fact and would not necessarily assist in the primary creative act. The models here show the relative frequency of step sizes, that give us, in Martin Gardner's apt expression, the mixture of pattern and surprise that is the root of pleasure in music.

Acknowledgements

I acknowledge the help of R.A.W. Bradford in writing this paper, and of Jim Backhouse of Manchester University who made the models.

References

[1] J. Beran. *Statistics in Musicology*. Chapman & Hall/CRC Press, Boca Raton, 2004, pp.32–34.

[2] K.J. Hsu & A.J. Hsu. "Fractal geometry of music." *Proc. Natl. Acad. Sci.,* Vol.87, 1990, pp.938–941.

[3] I. Holst I. *Tune.* Faber & Faber, 1962, pp.13–15.

[4] W. Pater W. *The Renaissance: Studies in Art and Literature.* Oxford, 2010, p.136.

[5] D. Parsons . *The Directory of Tunes and Musical Themes*. Spencer Brown & Co, 1975. Reprinted with additional material by G. Spencer-Brown, Bohmeier Verlag, 2002.

[6] D. Parsons. "Are composers copycats?" *New Scientist,* 24th March 1977, pp.688–689.