Is it too complicated yet?

September 19, 2013

Tim Blackwell, Goldsmiths, University of London

with Oded Ben-Tal (Kingston University)

Pattern = repetition + variation = regularity + randomness

But wherein lies complexity?

Reich's clapping

Steve Reich

clapping music for two performers (1972)

Directions for Performance

The number of repeats is fixed at 12 repeats per bar. The duration of the piece should be approximately 5 minutes. The second performer should keep his or her downbeat where it is written, on the first beat of each measure and not on the first beat of the group of three claps, so that the downbeat always falls on a new beat of the unchanging pattern. No other accents should be made. It is for this reason that a time signature of 6/4 or 12/8 is not given – to avoid metrical accents. To begin the piece one player may set the tempo by counting quietly; "one, two, three, four, five, six".

The choice of a particular clapping sound, i.e. with cupped or flat hands, is left up to the performers. Whichever timbre is chosen, both performers should try and get the same one so that their two parts will blend to produce one overall resulting pattern.

In a hall holding 200 people or more the clapping should be amplified with either a single omni-directional microphone for both performers, or two directional microphones; one for each performer. In either case the amplification should be mixed into mono and both parts fed equally to all loudspeakers. In smaller live rooms the piece may be performed without amplification. In either case the performers should perform while standing as close to one another as possible so as to hear each other well.

♩ = 160-184 Repeat each bar 12 times

162

Clapping *looks* simple (although could be perceived as highly detailed in performance).

The pattern is obvious - very repetitive, variation occurs through the phased superposition of the two parts.

How simple is the formal pattern?

The structure, without accents, can be described in a few dozen words.

A claps 111011010110 $\times$ 12 in 13 repeats. B claps the same pattern, beginning one beat displaced at each repeat.

It can also be described by a short program.

```
public static int[] clapping() {

    // 2 parts, 13 bars, 12 repeats of each bar, 12 beats per bar
    int[] onsets = new int[2 * 13 * 12 * 12];

    int[] pattern = { 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0 };
    for (int part = 0; part < 2; part++)
        for (int bar = 0; bar < 13; bar++)
            for (int rep = 0; rep < 12; rep++)
                for (int beat = 0; beat < 12; beat++)
                    onsets[beat + rep * 12 + bar * 144 + part * 1872] = part == 0 ? pattern[beat]
                            : pattern[(beat + bar) % 12];
    return onsets;
}
```

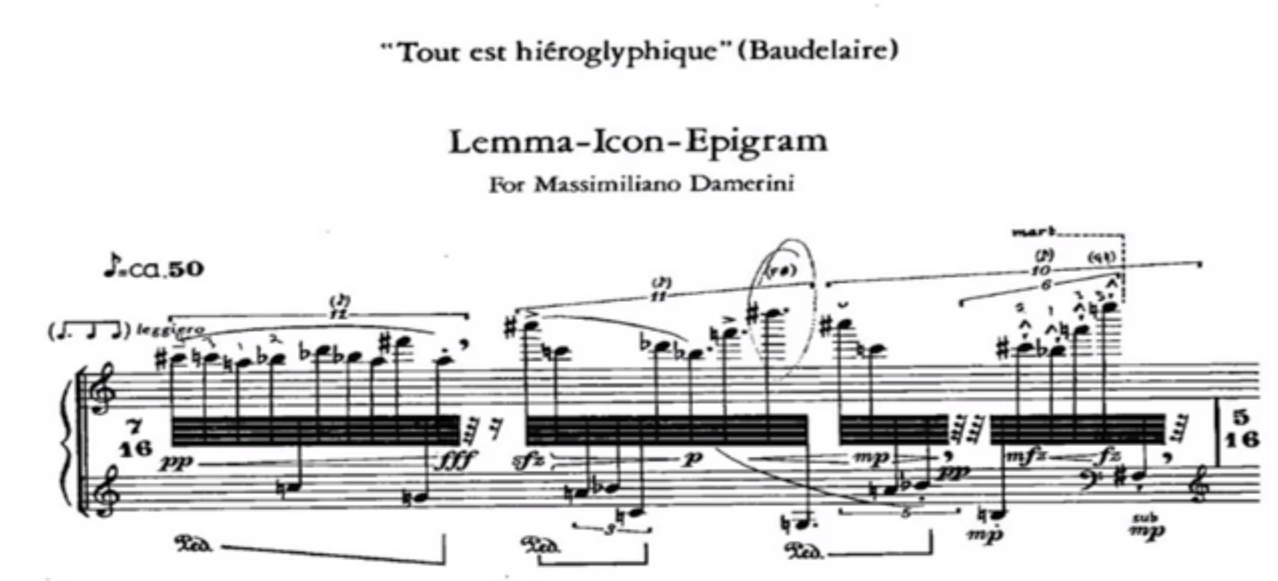
195 characters in shortened Java code

3744 digits in beat sequence 111011010110...

$\approx$ 340 characters in score (13 symbols in 13 bars, $+$ rep 12$\times$ instruction)

0.0710 dictionary words per symbol after LZ compression.

Ferneyhough's Lemma Icon Epigram



The first bar is split into 12 + 8 + 8 hemi-demi-semi quavers.

The first group of 12 is split into (11 hdsq + 1 hdsq rest) played in the time of 8 hdsq and then a semiquaver rest (= 4 hsdq rests), making 12 hdsq in total. This gives 11 notes of value 8/12 hdsq plus a rest of value 1/12 hdsq (together giving 8 hdsq) and a 4 hdsq rest.

Denoting a note a rest by $\langle\rangle$ and a fractional duration of an hdsq by $()$, then the first group is

$$11\left(\frac{8}{12}\right) \quad + \quad \langle 1\rangle\left(\frac{8}{12}\right) \quad + \quad \langle 4\rangle\left(\frac{1}{1}\right)$$

The second group has 11 hdsq played in the time of 8 hdsq.

$$2\left(\frac{8}{11}\right) \quad + \quad 3\left(\frac{2}{3}\right)\left(\frac{8}{11}\right) \quad + \quad 1\left(\frac{8}{11}\right) \quad + \quad 4\left(\frac{3}{2}\right)\left(\frac{8}{11}\right)$$

The final group has 4 hdsq plus a hdsq rest played in the time of 4 hdsq which itself is 4/10 of 8 hdsq. This is followed by a 8/10 hdsq rest and then 6 hdsq played in the time of 4 hdsq, played in the time of 4/10 of 8 hdsq. Finally a hdsq rest in the time of 10/8 hdsq.

$$
\begin{aligned}
4\left(\frac{4}{5}\right)\left(\frac{8}{10}\right) \quad &+ \quad \langle 1\rangle\left(\frac{4}{5}\right)\left(\frac{8}{10}\right) && + \quad \langle 1\rangle\left(\frac{8}{10}\right) \\
&+ \quad 6\left(\frac{4}{6}\right)\left(\frac{8}{10}\right) && + \quad \langle 1\rangle\left(\frac{8}{10}\right)
\end{aligned}
$$

In total, the rhythm of the first bar is

$$
\begin{aligned}
&11\left(\frac{8}{12}\right)+\langle 1\rangle\left(\frac{8}{12}\right)+\langle 4\rangle\left(\frac{1}{1}\right)+\\
&2\left(\frac{8}{11}\right)+3\left(\frac{16}{33}\right)+1\left(\frac{8}{11}\right)+4\left(\frac{24}{22}\right)+\\
&4\left(\frac{32}{50}\right)+\langle 1\rangle\left(\frac{32}{50}\right)+\langle 1\rangle\left(\frac{8}{10}\right)+6\left(\frac{32}{60}\right)+\langle 1\rangle\left(\frac{8}{10}\right)
\end{aligned}
$$

Simplifying (i.e. algebraically!),

$$
\begin{aligned}
&11\binom{2}{3}+\langle 1\rangle\binom{2}{3}+\langle 4\rangle\binom{1}{1}+\\
&2\binom{8}{11}+3\binom{16}{33}+1\binom{8}{11}+4\binom{12}{11}+\\
&4\binom{16}{25}+\langle 1\rangle\binom{16}{25}+\langle 1\rangle\binom{4}{5}+6\binom{8}{15}+\langle 1\rangle\binom{4}{5}
\end{aligned}
$$

Expressing in the lowest factor, namely $\frac{1}{825}$ of an hdsq,

$$
\begin{aligned}
&11\left(\frac{550}{825}\right)+\langle 1\rangle\left(\frac{550}{825}\right)+\langle 4\rangle\left(\frac{825}{825}\right)+\\
&2\left(\frac{600}{825}\right)+3\left(\frac{400}{825}\right)+1\left(\frac{600}{825}\right)+4\left(\frac{900}{825}\right)+\\
&4\left(\frac{528}{825}\right)+\langle 1\rangle\left(\frac{528}{825}\right)+\langle 1\rangle\left(\frac{660}{825}\right)+6\left(\frac{440}{825}\right)+\langle 1\rangle\left(\frac{660}{825}\right)
\end{aligned}
$$

```java
public static int[] ferneyhough() {
    int[] onsets = new int[28 * 825];
    int beat = 0;
    int end = 11 * 550;
    for (; beat < end; beat += 550)
        onsets[beat] = 1;
    onsets[beat] = 0;
    beat += 550;
    end = beat + 4 * 825;
    for (; beat < end; beat += 825)
        onsets[beat] = 0;
    end = beat + 2 * 600;
    for (; beat < end;  beat += 600)
        onsets[beat] = 1;
    end = beat + 3 * 400;
    for (; beat < end; beat += 400)
        onsets[beat] = 1;
    onsets[beat] = 1;
    beat += 600;
    end = beat + 4*900;
    for (; beat < end; beat += 900)
        onsets[beat] = 1;
    end = beat + 4 * 528;
    for (; beat < end;beat += 528)
        onsets[beat] = 1;
    onsets[beat] = 0;
    beat += 528;
    onsets[beat] = 0;
    beat += 660;
    end = beat + 6 * 440;
    for (; beat < end; beat += 440)
        onsets[beat] = 1;
    onsets[beat] = 0;
    beat += 660;
    return onsets;
}
```

23100 digits in beat sequence $1\overbrace{00\ldots0}^{824}1\overbrace{00\ldots0}^{824}\ldots$ where each beat is $\frac{1}{825}$ hdsq

334 characters in Java program (shortened version, but could do better by replacing some shorter loops with direct assignments)

About 60 characters needed to specify rhythm.. Of these, about 10 different symbols.

665 characters in English description

0.0102 dictionary words words per symbol after LZ compression

But the coding in $\frac{1}{825}$ hdsq is far removed from what we perceive. Each $\frac{1}{825}$ hdsq beat occupies

$$\frac{60}{400 * 825} = 1.818 \times 10^{-4} \text{ s} = 8 \text{ samples at 44100 Hz}$$

A performer needs to play to an accuracy of 8 samples, or 1/10 millisecond!

Use a more musical representation

hdsq beats prefaced by a time signature

$$
\begin{aligned}
&11\left(\frac{2}{3}\right) + \langle 1\rangle\left(\frac{2}{3}\right) + \langle 4\rangle\left(\frac{1}{1}\right) + \\
&2\left(\frac{8}{11}\right) + 3\left(\frac{16}{33}\right) + 1\left(\frac{8}{11}\right) + 4\left(\frac{12}{11}\right) + \\
&4\left(\frac{16}{25}\right) + \langle 1\rangle\left(\frac{16}{25}\right) + \langle 1\rangle\left(\frac{4}{5}\right) + 6\left(\frac{8}{15}\right) + \langle 1\rangle\left(\frac{4}{5}\right)
\end{aligned}
$$

codify in blocks (fractional duration in hdsq) beat pattern

In order to distinguish the time signature from the pattern, double up on the binary digits in the fraction and the pattern and use 01 as a separator:

$\frac{2}{3} = 11\ 00\ 01\ 11\ 11\ 01$

11 beats $= \overbrace{11\ldots1}^{22}$

$$
\begin{aligned}
11\binom{2}{3} &+ \langle 1\rangle\binom{2}{3} + \langle 4\rangle\binom{1}{1} \\
&= 2, 3, 11, 2, 3, \langle 1\rangle, 1, 1, \langle 4\rangle \\
&= 2, 3, 11\langle 1\rangle, 1, 1, \langle 4\rangle \\
&= 11\ 00\ 01\ 11\ 11\ 01 \\
&\quad 11\ 11\ 11\ 11\ 11\ 11\ 11\ 11\ 11\ 11\ 11\ 00\ 01 \\
&\quad 11\ 01\ 11\ 01 \\
&\quad 11\ 11\ 11\ 11
\end{aligned}
$$

Finally

$$2, 3, 11\langle 1\rangle,\ \ 1, 1, \langle 4\rangle,\ \ 8, 11, 2,\ \ 16, 33, 3,\ \ 8, 11, 1,\ \ 12, 11, 4,$$
$$16, 25, 4\langle 1\rangle,\ \ 4, 5, \langle 1\rangle,\ \ 8, 15, 6,\ \ 4, 5, \langle 1\rangle =$$

11000111110111111111111111111111110001
110111010000000001
11000000011100111101111101
1100000000011100000000110111111101
110000000111001111011101
111100000111001111011111111101
110000000001111100001101111111110001
11000001110011010001
1100000001111111110111111111111101
11000001110011010001

280 digits in sequence
0.207 dictionary words per symbol

(A decoding program should be added to the sequence)

The clapping beat sequence has a complexity of 0.0710 words per symbol.

Messiaen's Turangalila symphony

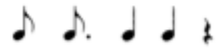
Hook proposes an algebra of rhythm to facilitate discussion of Maesiaen's approach to rhythm, especially in the context of the Turangalila symphony.

Hook argues that Messiaen's approach to rhythm was highly unusual (in the West, and for its time) and individual. He was influenced by Indian rhythms and to ametrical structures without regular repetition. He was also particularly attracted to rhythmic palindromes.

Some passages in Turangalila have 'intricate constructions involving many simultaneous rhythmic processes, each synchronised to its own clock, revolving in repetition; the resulting sensation has been described as an "image of constant uniformity and constant change" '.

Messiaen himself described some principles of his use of rhythm in *The Technique of My Musical Language* (1944): added values, exact and inexact augmentations and diminutions, rhythmic pedals and rhythmic superposition. However the descriptions remain simple and Turangalila was composed several years after publication of the book.

Hook attempts an algebraic descriptive language:



is represented, in units of semiquavers, by a sequence

$x = 2\ 3\ 4\ 4\ \langle 4 \rangle$.

In general $x = x_1\ x_2\ \ldots x_n$, $n = |x|$.

Retrograde $x_R = x_n\ x_{n-1} \ldots x_1$
$(2\ 3\ 4\ 4\ \langle 4 \rangle)_R = \langle 4 \rangle\ 4\ 4\ 3\ 2$

Augmentation, diminution $x \times q = q \times x_1\ q \times x_2 \ldots$.
$(2\ 3\ 4\ 4\ \langle 4 \rangle) \times 2 = 4\ 6\ 8\ 8\ \langle 8 \rangle$.

Concatenation $xy = x_1\ x_2 \ldots x_{|x|}\ y_1\ y_2 \ldots y_{|y|}$.

Ellision $x \circ y = \big(x_{|x|} == y_1\ ?\ x_1\ x_2 \ldots x_{|x|}\ y_2\ y_3 \ldots y_{|y|}\ :\ xy\big)$.
$2\ 3\ 4\ 4\ \langle 4 \rangle \circ (2\ 3\ 4\ 4\ \langle 4 \rangle)_R = 2\ 3\ 4\ 4\ \langle 4 \rangle\ 4\ 4\ 3\ 2$

Notation for repetitions: $x = p^q = \overbrace{p\ p\ p\ \ldots\ p}^{q \text{ times}}$.

The *composite* rhythm $x * y$ of two rhythms x and y is the rhythm heard when x and y are presented simultaneously. In the case that $\sum x_i \neq \sum y_i$, rests are padded onto the end of the shorter rhythm.

Hook isolates two sequences from Turangalila,

$$
\begin{aligned}
x &= 4\ 4\ 4\ 2\ 3\ 2 \\
&= (1^3 \times 4)(2\ 3\ 2) \\
y &= 4\ 4\ 4\ 2\ 3\ 2\ 2\ 2\ 2\ 3\ 3\ 3\ 1\ 2\ 3\ 4\ 8 \\
&= x\ (1^3 \times\ 2)(1^3 \times 3)(1\ 2\ 3\ 4\ 8)
\end{aligned}
$$

Hook observes that many of Messiaen's rhythms grow from initial segments (seeds) by a systematic process.

Indefinite repetition $X = \bar{x} = x\ x\ x\ \ldots$. The period of repetition is $|x|$. The period might not align with the bar lines and simultaneous repetitions might have different periods. They might eventually abruptly end, being cut-off mid-cycle or even in mid-duration.

Expanding, contracting progression A seed of a single duration is incremented or decremented (chromatically if the increment/decrement is 1). For example

$$X = 7\ 8\ 9\ 10\ 11$$
$$Y = 48\ 47\ 46\ 45\ 44$$

Complex generation (i)The seeds might have more than one interval (ii) the increments might vary cyclically (iii) the rules change at each iteration.

The cyclic increment occurs in $\langle 2\rangle 3\ 5\ 8\ 9\ 11\ 14\ 15\ 17$, or $\langle 2\rangle\ 3(=2+1)\ 5(=3+2)\ 8(=5+3)\ 9(=8+1)\ldots$, i.e $x_{i+1} = x_i + 1 + (i \mod 3)$.

An example of expansion of a more complex seed is

$(1\ 4\ 7\ 6\ 5\ 3\ 2)(8\ 11\ 14\ 13\ 12\ 10\ 9)$.

Hook also identifies

Recursion, pattern shifting, interruptions, irregularities and relationship to pitch structures

At rehearsal number [14] in the fourth movement of *Turangalîla,* a single wood block plays rhythm *IVb,* Messiaen's favorite periodic rhythm, whose period is 52 𝅘𝅥𝅯. Meanwhile the small Turkish cymbal and vibraphone play *IVa,* the simple chromatic contraction-expansion rhythm of period 240 𝅘𝅥𝅯. The snare drum plays *IVd,* a complex recursively generated rhythm with period 132 𝅘𝅥𝅯, and the double basses are playing chromatic scales to rhythm *IVc,* a rhythm of period 84 𝅘𝅥𝅯 whose seed consists of four segments in progressive augmentation. At the same time, the piccolo, bassoon, and *ondes martenot* (soon to be joined by trumpets) present the scherzo theme from the beginning of the movement. The celeste and glockenspiel are playing a rhythmically simplified version of a birdsong introduced previously by the piano; the piano is now engaged in new decorative figuration all its own. In the fifth measure the themes from the two "trio" sections heard earlier are added to the mix, the first in the winds and the second in the strings. Eleven measures later the trombones enter in grand fashion with the "statue" theme, bringing the number of independent parts to ten.[31]

How is a listener able to process such complexity? How can so many simultaneous layers of sound be perceived with any degree of clarity? How can the lonely rhythm played by the wood block (which, incidentally, is marked *piano*, while many other instruments are playing *forte* and *fortissimo*) even be *heard*, much less *understood*? In Messiaen's words, listeners to his music

Figure 1: Julian Hook, 'Rhythm in the Music of Messiaen: An Algebraic Study and an Application in the Turangalla Symphony', Music Theory Spectrum, Vol. 20, No. 1 (Spring, 1998), pp. 97-120

[Listeners] will be responsive to [complexity] the day their ears are accustomed to it. It's not essential for listeners to be able to detect precisely all the rhythmic procedures of the music they hear, just as they don't need to figure out all the chords of classical music. That's reserved for harmony professors and professional composers. The moment [listeners] receive a shock, realise that it's beautiful, that the music touches them, the goal is achieved!
Messiaen, Music and Color, 83.

Measuring complexity

The fundamental insight is that the description length (in Java, English or in dictionary words) sets a scale from simplicity (repetition: repeat 1111 32 times) to pure variation (a sequence of 128 random 1's and 0's that itself occupies about 128 characters.)

Complexity is assumed to lie between these extremes; a regime of balanced repetition and variation.

Complexity can be classified into the complexity of the artefact (or of the ensemble of which artefact is typical), or the complexity of the process producing that artefact (or of representative from the corresponding ensemble). A third class which intersects the previous is the degree of organisation.

1. Difficulty of description Information, randomness, entropy, minimum description length, Kolomogorov complexity, Code length, fractal dimension, compressibility.
2. Difficulty of creation. Computational complexity (time and/or space), logical depth.
3. Degree of organisation Fractal dimension, excess entropy, statistical complexity.

Seth Lloyd, 'Measures of complexity: a non-exhaustive list'. Control systems, vol 4, pp 7 - 8, 2001.

Lloyd lists over 40 measures and considers (and that was in 2001) the list to be incomplete.

Another fundamental divide is between statistical measures and those that consider a single instance.

A piece of music, as with a work of art, is a single artefact. However are there broad categories that we slot the artefacts into? Are there many versions of clapping, each different in detail but employing the same algorithmic (generative) scheme, and does each version seem roughly as complex as any other version?

Insight from psychology

Berlyne: hedonic value is a sum of reward and pleasure. High arousal is aversive (negative hedonic value) and any stimulus that reduces arousal will be rewarding and pleasant. Any stimulus that produces a moderate increase in arousal will be rewarding and pleasant.

When a complex stimulus becomes more familiar and less novel, the observer moves from C to B. Something that is high in novelty and low in complexity will have medium arousal (region B) and progressive loss of novelty will invoke a move to region A.

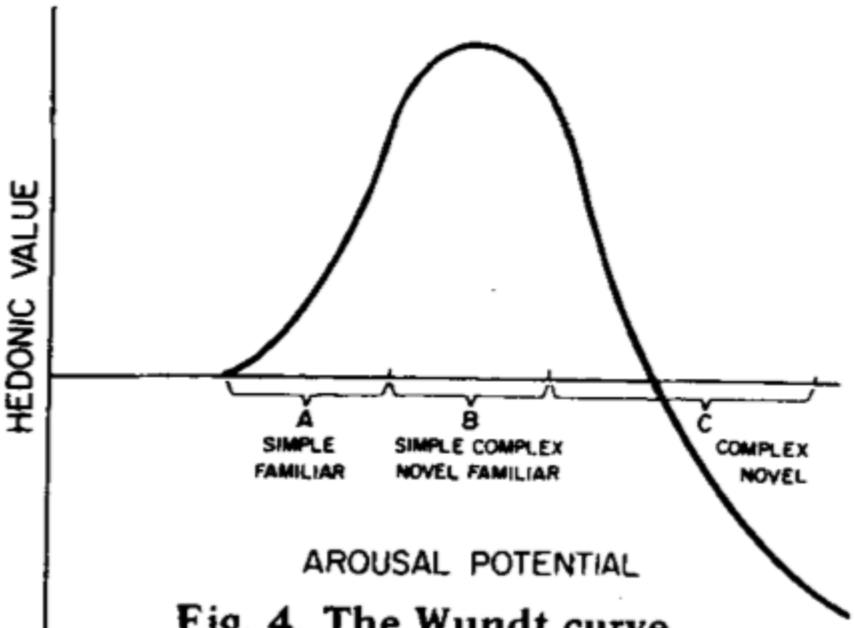


Fig. 4. The Wundt curve.

This is often simplified to a camel hump between simple (region A) and complex (region C) with a peak in the middle representing maximum pleasure (or interest).

This is all in the eyes (and ears) of the observer. Do objects have an intrinsic complexity? Can this be measured? Does it correspond to perceived complexity? And of so, does this correspond to our interest/pleasure in a piece of music? And does this interest change in time?

Complexity science

Many systems in the physical, biological, social and mathematical world appear to have something in common: they are 'complex'. This vaguely means that they appear to have properties beyond the sum of their parts, they show emergent behaviour and typically self organisation.

(Or, more conservatively, analysis based on weak interactions between constituent components fails. The systems often exhibit non-linearity which, even in systems with a few components, can lead to intractability.)

Complexity measures - requirements

How can we measure something that is not defined? It's like measuring the 'biology' in something.

According to the current creed a qualitative measure of complexity would be zero at the limits of order and randomness and with a camel hump between them - the inverted U.

A unique measure that can characterise many systems from different domains?

The measure has one hump? This is a very poor requirement for a property since many functions are hump shaped.

However a definition of complexity might follow the adoption of a measure.

There already exists measures of randomness (Shannon and Kolmogorov). The fact that they do not have a hump doesn't matter, they still form a measure. For example complexity might exist for Kolmogorov complexity in range 0.3 - 0.5.

Some thoughts

Complexity is something between order and randomness.

Something that is repetitive and entirely predictable holds little interest. It is perfectly understood.

Something that appears entirely random and defies our efforts to understand is also ultimately boring.

In fact the random becomes repetitive in the sense that the statistical qualities (as perceived)become predictable.

So the entirely predictable and the seemingly random, although opposites, become equally unsatisfying.

A complexity measure would be very useful. Two applications that spring to mind are evolutionary music and live algorithms.

The fitness bottleneck makes user fitness evaluation problematical. A measure could act as a tuneable fitness function.

Complexity, or better still a measure of 'information' might serve as a machine aesthetic. This would enable live algorithms to assess their own patterning, and they would not have to base their decisions on what we might want from them.

What might alien music look or even sound like? We might have to reduce it to a suitable representation and consider alien music in terms of information streams.

Huge application scope for automatic accompaniment of games, music composition for a target market etc. etc.

Although it seems very unlikely that music, which lives in both cognitive and social domains, can be expressed by a single number, perhaps some aspects of music can.

We hope for a principle for the generation and the evaluation of computer music.

Kolomogorov complexity

Better thought of as a measure of randomness.

The length (in bits) of the shortest program that would generate the object in question. The underlying assumption is that all objects can be represented by strings s of symbols, and indeed by binary strings. Technically the K complexity is the shortest encoding of the Turing machine that outputs s and stops.

The program is clearly a description of s and the shortest program provides a minimal description.

Suppose that a string s has 1000000 symbols, each either A or B.

Then $s = AAA \ldots AAA$ can be produced by the minimal program

```
repeat 1000000 times: print A
```

and another string $s' = AABBABABABBBAAABABBBAB \ldots$ where each symbol is random cannot be produced by any program shorter than

```
print ``AABBABABABBBAAABABBBAB...''
```

Although s has 1000000 symbols, it has K complexity of about 30 (in this encoding); s' however has a K complexity of about 1000000.

K complexity is not computable, but a series of upper bounds is computable.

Compressibility provides an upper bound and is easily computable using the LZ78 algorithm:

```
w = ""
while(more input)
   s = next symbol
   if ws in dictionary
      w = ws
   else
      add ws to dictionary
      w = " "
```

Li and Vitanyi have suggested the information distance as a substitute for the unknowable K complexity. The information distance between two strings x, y is the length of the shortest program that computes x given y or vice versa. It is equal to $\max\{K(y|x), K(x|y)\}$

It is also non-computable, and not normalised. But the normalised information distance

$$d(x, y) = \frac{\max\{K(x|y), K(y|x)\}}{\max\{K(x), K(y)\}}$$

varies between 0 and 1 for strings of any length . The K complexities can be estimated by the length of the compressed string.

Shannon entropy

This is another randomness or entropy measure, but this time it is statistical in nature.

An information source produces a stream of symbols. Each symbol is a discrete random variable with an associated probability distribution $p(x)$.

The information associated with each symbol is $\log 1/p = -log\ p$. Information relates directly to the minimum description length of messages: use the more probable symbols for the more probable events. The logarithm ensures that information is additive (the underling probabilities of independent events multiply together).

Alternatively, the surprise of receiving a symbol is the inverse probability of that symbol.

Suppose we want to report on a search. Then A might encode ‘Not found, not found’ and ’B’ might encode ‘I’ve found it!’. We send AAAAAAAAAAAAAAAAAAAAB. B is very improbable, but has a significant interpretation. We might try a coding where any symbol other than A is insignificant, but all symbols are equally probable, but then we would need a huge alphabet.

The entropy of a sequence is

$$H = -\sum_i p_i \log(p_i)$$

where p_i is the probability of the ith symbol. The entropy rate $h = \lim_{L\to\infty} H/L$ is an extensive quantity (L = sequence length) and ranges from 0 (complete order) to 1 (all symbols equally likely).

However H and h are not sensitive to the order of symbols and hence are of little use of complexity measures. A quantity that has been discovered several times and has several different names, does however give an indication of complexity: the excess entropy.

The excess entropy is the subextensive part of the entropy: $\lim_{L\to\infty} H(L) \sim E + hL$.

Suppose a sequence is periodic. Then for any L larger than the period, $H(L) = 0$ $(= h)$ since we can replace the repeating subsequence by a single symbol. However for L less than the period, $H > 0$ since as each symbol is revealed we do not yet where we are in the subsequence and hence each symbol carries an uncertainty (and hence a finite entropy).

A similar argument can be made for non-periodic information sources, except that $h > 0$; E is how much more uncertain a sequence is compared to an unstructured sequence that is composed of the same symbols occurring with the same probability. This excess of entropy is due to *structure* i.e. correlations between the symbols.

There are many other measures in the literature that relate to E and to the mutual information between subsequences, $I(X;Y) = H(X) - H(X|Y)$.

For example, the recent predictive information rate of Abdallah and Plumley,

$$b = I(X_t; \overrightarrow{X}_t | \overleftarrow{X}_t)$$

which can be interpreted as the mutual information between the present symbol X_t and the future ($\overrightarrow{X}_t$) given the past ($\overleftarrow{X}_t$).

In fact $b = H(\overrightarrow{X}_t | \overleftarrow{X}_t) - H(\overrightarrow{X}_t | X_t, \overleftarrow{X}_t)$ which is small for random sequences where each symbol is independent of each other (both terms are equal), or for regular (constant or periodic) sequences (the terms are equal). It seems to possess the expected attributes of a complexity measure.

The redundancy or multi-information rate

$$\rho = I(\overleftarrow{X}_t; X_t) = H(X_t) - H(X_t|\overleftarrow{X})$$

measures the reduction in uncertainty of the current symbol once the past is taken into account. Abdullah et al 2012 define a 2D information space spanned by h and ρ. It turns out that b peaks at intermediate values of h and ρ.

They then constructed their melody triangle. A user could navigate through information space, listening to melodies that are generated according to the co-ordinates. They conducted preliminary trials with a very small group of people and recorded which areas of information space were frequented most often.

Unfortunately it seemed that subjects steered clear of the centre of the space - high b, complex melodies - and aimed for a corner corresponding to repetitive melodies of just one note!

Other measures

Logical depth: The runtime (or similar resource) of the smallest program (or amount of time needed for the formational processes to complete). Some qualification is needed since a slightly longer program might run in much less time.

Suppose a string solves a problem that is quite simple to pose, and the string itself is quite short. If it takes a long time to compute this string, the string is 'deep'.

Schmidhuber's theory of aesthetics

Beauty is directly related to an observer's ability to compress i.e. to the description length the artefact, based on the observer's compressor, which itself might change in time. The compressor represents the observer's past experience.

No matter how beautiful something is, repeated exposure may render it boring. Interestingness is the rate of change of beauty. When previously random parts of the artefact are compressed by the observer (as the observer's algorithm improves), the subjective beauty increases, and as long as this process continues, we remain interested. Remaining interested is the driving force, both of creators, and of the observers.

Schmidhuber's definition is really just a computational model of Wundt/Berlyne. In particular, the model provides a mechanism for the change of interest over time - movement from right to left along the Wundt curve.

But some music is appealing time after time, just as a favourite meal continues to satisfy us even of we can predict the taste. In fact the mere prediction of the taste can cause us to salivate! Some aspects of our need for music are similar to appetite.

Sounds have have an emotional quality (a scream, a sob. . .). Can we become inured to an emotional message?

Incompressible music - or at least music that cannot be compressed any further - can still allow our minds to wonder, to freely associate. An underlying structure might be present, but also an intrinsic ambiguity and lack of clarity provokes speculation.

We are not trying to compress the work any further, we are happy with our current level of comprehension; we enjoy the springboard into the unknown, the opportunity to form juxtapositions in other (not directly related to the work) domains.

Anticipated problems

Measures may take a large sequence to stabilise on a value - but we might react to a piece of music after just several events.

Symbol streams are timeless, yet the temporal relationship of events is important to human appreciation of music. Music is not a static pattern, but is a moving pattern.

Correlations between events that are distantly separated in time may be lost to us because of our limited memory and concentration, and patterns that quickly fly past may escape our attention.

To some extent a symbol sequence does have a near and a far time - near and far symbols, and this may affect any finite calculation of complexity since a large separation symbol group may not occur often enough in a finite sample to count much in the calculation.

However just one or two repetitions of a theme might strike us as significant and be very memorable, although their contribution to the complexity measure could be small.

We are also tolerant of imperfections in the pattern.

AAAABBBAAABAABBBAAABAAAAABBBAAAAAABAAA

might be heard as

...BBB...B...BBB...B...BBB...B

i.e. as

XBBBXBXBBBXBXBBBXB

(On the other hand the observation that the X's are of uneven duration might strike us as delightful, as interesting, and this would be captured by a complexity measure such as compressibility.)

We focus on the pattern - BBBXBY - and ignore the randomness represented by X and Y. We manage to pick out a pattern; it has more significance than the background randomness.

If the string is very long, then the measures will find this pattern since eventually there will be repetition. The dictionary will contain words BX, BY, BBBZ etc where X, Y and Z are all different, and eventually the dictionary will contain BXBBBY for all occurring X and Y.

Empirical studies

Povel and Essens 1985. Reproducibility of temporal patterns. Subjects had to listen to the patterns and then tap them out. Thul and Tossaint 2008 derive a measure of performance reproducibility based on the results - the rhythm performance complexity.

Shmulevich and Povel 2000, using the same data set as PE 1985, compiled a table of perceptual complexity based on an empirical study with musicians.

Essens 2005. A repeat of PE 1985 but with a different set of patterns this time. Performance and perceptual complexity.

Fitch and Rosenfeld 2007. A beat tracking study. A pattern set of various syncopations.

Thul 2008, masters thesis. Compiled tables of the above studies. They took a large number of complexity measures, many of which have been deliberately designed for measuring rhythmic complexity, and compared the calculated complexities with the empirical results.

None of the measures were found to reflect the difficulty humans have in performing rhythm patterns, but they did manage to predict to some extent how well people recognise a rhythmic metrical structure.

Thul and Toussaint 2008, in an extension to Thul's work, found that measures based on statistical properties of onset intervals were less reliable measures of performance complexity and that mathematical measures of perceptual complexity performed better than performance complexity for the purpose of predicting human perception complexity.

Thul concludes that half of the 55 complexity measures correlated quite well ($r_s^* > 0.5$) with the empirical measure of complexity. There appears to be evidence that complexity measures are reflecting human perception of rhythm complexity.

(Apart from rhythm studies, it is worth noting Streich's 2006 PhD thesis where complexity measures for audio were developed. He asked subjects to rate music and a comparison with the measures was made, but no strong correlation was found.)

In summary, with regard to the four empirical studies, domain specific measures are better predictors, there is evidence that, in general, theoretical measures are capturing salient features of music complexity. The better performing seem to be those that have been hand crafted to so the job.

However the studies are limited and feature isolated rhythm patterns rather than actual music.

In addition, the use of music derived measures, is questionable for reasons of circularity.

Is it complicated enough yet?

Repeat experiments but using more realistic exemplars.

The Hook algebra is a way of generating patterns and is based on music practice.

Production rules in general need infinite memory - the excess entropy is infinite. Would need non information based measures, but K complexity based measures are non-computable.

Interesting music might lie closer to order than to randomness. Different more sensitive measures needed to magnify this regime.

The problem of time.

The problem of human judgement based on quite short excerpts.

The problem of human patterning from imperfect information.

Thank you for listening. On a scale of 1 to 10, how complicated was it for you?