Stylistic randomness: about composing *NTrope Suite*

SHLOMO DUBNOV

Department of Communication Systems Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
E-mail: dubnov@bgumail.bgu.ac.il

This paper describes the ideas and methods that led to the writing of an algorithmic composition *NTrope Suite*. This piece, for solo recorders and voice, was generated by 'mixing' works by different composers from different eras. The idea behind the work was to examine random generation procedures that could maintain stylistic properties typical to the reference works. The interesting property of this method is that it implements a sort of 'statistical learning', that optimally preserves the properties of the reference pieces and also properly 'generalises' them so as to create a new valid work. The musical result is very coherent, maintaining both stylistic resemblance to the reference music and exhibiting some surprising originality as well. Theoretically, the resulting piece is closest to the reference works in terms of mutual entropy. The algorithm and its theoretical significance are discussed in the paper.

1. INTRODUCTION

This paper describes the ideas and methods that were employed for composing *NTrope Suite*, a short work (approx. 8 min) consisting of three musical movements for solo recorder and one text movement, all created by the computer. The piece was composed at IRCAM, Paris, and premiered at the Biennale for Contemporary Music in Tel-Aviv in 1998.

There were no composition algorithms employed in the usual sense of music programming that represent musical knowledge and formulate the composition in algorithmic terms. Rather, the whole process of composing *NTrope* was based upon a random recombination of other musical works in a nontrivial manner. Thus, the result could be considered as a completely aleatory process of 'deconstructing' (and of course reconstructing) other musical works. The most surprising effect of the piece, in my opinion, is that the resulting composition is very coherent. The coherence is in the sense that no 'cut and paste' occurs and the melody flow is never disrupted by the fact that its every segment is drawn from another musical work. This is not a collage work, since the resulting piece is not heard as a juxtaposition of musical quotations. This is not a traditional algorithmic work either, since no 'formal' compositional ideas were specified ahead of time.

2. DESCRIPTION OF THE ALGORITHM

Let us start with two different works, either by the same composer or by two different composers. We shall refer to the two works as our 'learning samples' A and B. Then let us choose one of the samples A or B and randomly draw a single note from it. We shall denote this first note by 'a'. The next note to be composed is again randomly drawn either from A or B, but in a more 'structured' manner. The possible choices for note two are now all the notes in a chosen work that have 'a' as their predecessor. This means that we look for all pairs of notes that have 'a' as the left member. Let us say that we find three such pairs: {ab, ac, ab}. Then we randomly draw one of the pairs (actually we need only the right note since we know that 'a' is on the left) and put the second note in our new piece. Note that in our example the probability of selecting 'c' is 1/3 and that of selecting 'b' is 2/3. Let us assume that we choose 'b' and we have the two notes 'ab' in our piece. Then we apply the same process to compose the third note, fourth note and so on. What is important to note is that our procedure is designed so that the 'memory' always grows (at least until no patterns can be found in the reference piece). To explain this point let us proceed to compose the third note. In order to do so we use the two composed notes 'ab' as our memory, i.e. the next selection will be 'in context' of the existing notes. So we repeat the game of selecting a reference piece A or B and then look for all triplets of notes that have 'ab' on their left side in that piece. Naturally, we cannot always find the whole context. Maybe for the first two, three, or even ten or fifteen notes, there will be an exact match in the reference piece, but obviously that algorithm will not be able to match very quickly the whole past of the new piece segments in the works A or B. (Otherwise it would be a complete recopying. Such a situation might occur only if the works A or B are the same). So, here comes the tricky part of the algorithm: In order to compose the next notes, we actually search for the places in the reference piece that match the longest suffix in our new composition. What we mean by suffix is the longest segment from the last composed note backwards (in other words, the longest 'tail') that matches a sequence of notes in the reference piece. Naturally, the reference piece (from which the next note candidates will be drawn by the above procedure) is randomly chosen at first. If we end up with a single match, the next note will be simply copied. If there are several candidates that

https://doi.org/10.1017/S1355771899002046 Published online by Cambridge University Press

have the same longest suffix, one of these appearances is chosen at random and then the next note that follows the selected sequence is copied from the reference piece to our composition.

So far the whole process is trivial. Basically, it is no more than randomly recombining segments from the two reference works A and B, taking precise care that the next note best matches (or best continues) the music that was written so far. This is why the transitions are so smooth. Actually, there is almost no way of telling that the recopying process has jumped from one reference piece to the other. Whenever the jump happens, the algorithm will find the longest match and thus produce the best continuation.

Naturally, the process might fail if there are no common notes between the two pieces. In order to avoid this, we use not the actual notes but rather the intervals. Now, lack of continuation will occur only if the two reference pieces (composers) have no common motives or melody segments. From our experience, this does not occur practically.

3. STATISTICAL LEARNING ASPECTS

The interesting aspect of our method can be seen if the whole process is considered as a sort of statistical learning algorithm. By statistical learning we mean a method of acquiring certain statistical properties of a data source so that new sequences can be created, having the same properties as the source. At first it might seem strange to acknowledge the relation between our process and statistical learning, since normally one considers learning as some sort of information extraction, linked to some compact knowledge representation. In the process of learning, one usually finds out what is 'essential' in the data so that 'understanding' might occur. That is why simple recopying contradicts learning, since it does not compact things and seems not to be able to extract the essential from the irrelevant. Nevertheless, our method is in fact a learning process, as we shall explain below.

One of the main purposes of learning is creating a capability to sensibly generalise. Composers are interested in finding out the possibilities of certain musical material, without necessarily 'explaining' it. So, let us consider here the possibilities that a piece (or a set of pieces in a given style) offers. Statistical analysis of a corpus reveals some of the recombination possibilities that comply to constrains or redundancies typical of the particular style. The concept of redundancy is closely related to information or entropy.

4. INFORMATION AND ENTROPY

As Claude Shannon, the father of information theory, expressed in his work, *The Mathematical Theory of Communication* (Shannon 1949), the concept of information concerns the possibilities of a message rather than its content: 'That is, information is a measure of one's freedom of choice when one selects a message.' The measure of information is done by means of entropy, sometimes also called 'uncertainty'. Let us assume that we have a sequence of notes (a melody). If the notes are generated in an independent manner (such as in John Cage's 'Incidental Music' or his I-Ching technique), the best description of that music could be in terms of frequencies of appearance of different notes. Shannon suggested that if we look at large blocks of symbols (long sequences), this information 'reveals itself' without any need to calculate probabilities. This property (called AEP[1]) basically says that if we look at a long enough sequence, we will see only the 'typical' messages. The entropy of the source defines the amount of the typical messages. That is why a completely random (aleatoric) sequence, where all the notes appear with equal probability, has maximal entropy: the number of typical sequences is equal to all possible sequences. On the other hand, the more structure a sequence has, the less will be the number of typical sequences. The exponential rate of growth of the number of typical sequences, relative to the number of all possible sequences of the same length, is the entropy. Thus, looking at longer blocks of symbols we approach the 'true' statistics of the source.

Shannon played his famous game of generating random English text. In this connection, he examined the frequency of word correlations in the English language. Pairs of words, which often appear together, show a higher degree of redundancy than less common pairs. Shannon showed that a randomly generated string of words could sound remarkably like meaningful English, so long as each word had a high correlation with the word before it. The resemblance to English is even greater if the nonsense string is generated using word triplets, rather than word pairs.

Shannon actually achieved an effect like this without a computer. Taking a novel, he picked a first word at random, then found the next place in the novel where this word appeared. He added the following word to the generated text, then repeated the process until he had a fair-sized sentence such as:

> The head and in frontal attack on an English writer that the character of this point is therefore another method for the letters that the time of who ever told the problem for an unexpected.

This word information in communication theory relates not so much to what you do say, as to what you could say. Presumably, the mimicry would be even more accurate if we analysed word triplets

[1]Asymptotic Equipartition Property.

https://doi.org/10.1017/S1355771899002046 Published online by Cambridge University Press

instead of just word pairs. This would take up a good deal more computing time.

Using Shannon's idea, several curious applications appeared that attempt to 'rewrite' a given text in a style of some famous writer or change the dialect of the English language to several optional dialects. Here we describe two such examples.

The Dialectiser, *by Samuel Stoddard*
http://rinkworks.com/dialect/
The translation is done via a huge series of search-and-replace operations. The translation programs do not analyse sentence structure or keep track of parts of speech. They merely look for certain sequences of characters and replace them with other sequences. For example, in the Redneck dialect, occurrences of ing are replaced with in', and occurrences of potato are replaced with patootee, and so on. The search strings are sometimes whole words, sometimes parts of words, or sometimes more than one word.

The Shannoniser, *by Josh Senyak*
http://www.nightgarden.com/shannon.htm
This program analyses the style of the document, using word pairs. If you have selected an editor, the script combines the editor's style with the style of your document. (These 'preset' styles come from word-pair analysis of a virtual 'board of editors'.) The script now generates a piece of random text, using the desired style.

To demonstrate how the programs operate on a given text, we activated the programs on the abstract paragraph at the beginning of this paper. Here are the results:

Abstract 'Shannonised' in the style of Lewis Carol:

> Abstract: This paper describes the claws that grand? Lend her hands. Beware the French for fiddle-de-dee? Lend her your little boy, and burbled as well. The interesting property of this method implements a right-hand foot into mutton pies, for solo recorders and properly generalises them to writing of flame, the original works in uffish thought he knows it teases. Twas brillig, cried the Jabberwock, that bite, and gimble in terms of some surprising originality as in his joy. Theoretically, and shun the melodic style of flame, for solo recorders and shun the original works in uffish thought he sneezes, and methods that maintain statistical learning, that maintain statistical properties of the work. I weep. Come to be beheaded!, the resulting piece, with its head. Twas brillig, was to cry. This paper describes the work. The guinea-pigs cheered. You sha'n't be beheaded!

. . . and the same abstract dialectised by the Muppets' Swedish Chef:

> Thees peper descreebes zee idees und methuds thet led tu vreeting ooff un elgureethmic cumpuseeshun Ntrupe-a Sooeete-a. Thees peeece-a, fur sulu recurders und fueece-a, ves genereted by 'meexing' ooff vurks by deeffferent cumpusers frum deeffff-erent ires. Um gesh dee bork, bork! Zee idea beheend zee vurk ves tu ixemeene-a rundum genereshun prucedoores thet meeentein steteesticel pruperteees typeecel tu zee meludeec style-a ooff sume-a refference-a vurks. Um gesh dee bork, bork! Zee interesteeng pruperty ooff thees methud implements a surt ooff steteesticel leerneeng, thet oopteemelly preserfes zee prupertees ooff zee refference-a peeeces und pruperly genereleezes zeem tu creete-a a noo cuherent vurk. Zee moosee-cel resoolt is fery cuherent, meeenteining but a styleestic resemblunce-a tu zee ooreeginel vurks und ixheebiting sume-a soorpreesing ooreeginelity es vell. Zeeureteecelly, zee resoolteeng peeece-a is clusest tu zee ooreeginel vurks in terms ooff mootooel intrupy. Bork bork bork!

It is interesting to note the basic difference between the two approaches. While Shannoniser imitates by copying parts of sentences from editors' works, the dielectiser works more on a morphological level, changing letter combinations so as to produce a funny pronunciation. Although no control over content (meaning) of the text can be exercised using the above methods, the two applications exhibit the idea that certain stylistic properties of a text can be varied in interesting ways.

5. MARKOV SOURCES

Before finally arriving back at our *NTrope* method, which, at least in terms of the algorithm itself, obviously differs from Shannon's procedure above, we need to say a few words about Markov sources. By Markov sources we mean processes that produce new symbols depending on a certain number of past symbols. In other words, the whole process has some kind of memory whose effect is that new symbols are generated in a manner dependent on this memory. Old symbols influence new symbols by determining the probability of generating certain continuations. Markov sources are of different orders. Simple independent random sequences are considered as Markov of order zero, then processes of a memory of a single symbol are Markov of order one, of a memory of length two are Markov of order two, and so on.

The nice property of Markov sources is that they also fulfil the AEP property. What this means is that by considering large blocks of symbols (i.e. looking at long sequences), the statistics of the Markovian

https://doi.org/10.1017/S1355771899002046 Published online by Cambridge University Press

source is revealed as well. Actually, Shannon's word pair procedure captures the first-order Markov property of the English language. In music, Markov processes were investigated by Hiller (1959), Xenakis (1971), Ames (1989) and others. This is contrasted to style understanding and style 'imitation', where the latter usually relies on knowledge-based approaches, i.e. building of complex models that describe the rules of a specific composer, style or era (Cope 1991).

One of the main problems with the application of Markov theory 'as is' is that the true order of the process is unknown. Our algorithm tries to eliminate the main difficulty behind Markov modelling, i.e. handling large orders, or even more interestingly, handling sources of variable-length memory. In music it is reasonable to assume that the order is variable since we sometimes use short melodic figuration, sometimes longer motives, and also rather long melodic phrases. Here we come to the main claim of the *NTrope* procedure:

Claim:
NTtrope procedure creates a new sequence whose entropy is a mix of the entropies of the sources from which it was derived.

Remark:
When the learning examples are all drawn from the same musical corpus, we arrive at a situation where the resulting piece is equivalent to the entropy of the source. The new sequence could be considered as a 'legitimate' exemplar of the same statistical source that describes (creates) all the learning examples. Musically speaking, this means that if the corpus contains works by the same composer or compositions in a particular style, the new piece will be, at least in information theoretic terms, another piece in the same style. In the case when our learning exemplars belong to different sources, even more interesting possibilities can occur. This situation is known as the 'two source problem' (El-Yaniv, Fine and Tishby 1997) in statistics. Thus, the claim of our procedure is that the resulting random sequence is a realisation of a hypothetical source whose statistical similarity to the two original sources is maximal. We will not prove the claim in a rigorous manner here. Below we provide an Informal Proof:

Informal Proof:
The probability of a long sequence can be calculated from conditional probabilities for each symbol. Taking the probability of the first symbol, we multiply it by the conditional probability of the second symbol given the first, then multiply again by the conditional probability of the third symbol given the first two, and so on. The complete process will give a total probability of a sequence.

Now, in order to find the probability of the mixture we need to show that the conditional probability of generating a new symbol given some existing sequence is the correct combination of the conditional probabilities of the two sources. Our procedure does exactly this thing: by jumping between the reference works we have an *a priori* probability of using one of the pieces (let us say that this probability is 1/2 so the mixture is in equal proportions). Once the choice of selecting the reference work is done, we look for the set of longest suffixes in our new piece that exist in the reference piece (see the description of the algorithm above). The continuation note will be drawn randomly from this set. This search-and-draw process performed upon the candidate sequences is equivalent to drawing a continuation from a source whose distribution is an empirical conditional distribution of our reference piece. In other words, instead of creating big tables of conditional distributions (Markov tables) for all possible sequences and their continuations, we simply search each time for all the candidates for continuation in the reference work. If the reference piece is long enough, this is a good approximation to the actual conditional probability that represents the reference work.

So, to sum up, the first step of the mix (choosing the reference source) determines the amount of influence of each source (say 1/2). Then, the suffix search gives the empirical conditional probability of that source, as required. It can be shown mathematically that the resulting process is maximally similar (maximal mutual information) to the stochastic models that represent the two reference works (El-Yaniv *et al*. 1997).

What is essential for understanding the 'similarity' idea is that our mixing process creates a new sequence whose probability is an average of the probabilities of the sources. Since, as we have already mentioned above, the behaviour (probability) of a source reveals itself when we look at long sequences, the typical sequences in the new piece are a mixture of typical sequences from the two sources.

6. SOME MUSIC EXCERPTS

Let us now analyse some excerpts from the *NTrope Suite*. Each movement in the piece is composed from a mixture of two pieces, each belonging to a different composer. The interpolated movements are: Purcell–Grieg, Satie–Handel, Interpolated Text, Mozart–Scriabin.

The figure presents the first ten bars of the Purcell–Grieg Mix. Both the beginning melodic line and the harmonic structure seem to be reminiscent of Grieg. The third movement is an interpolated text that was derived using the same methods. It uses two reference

https://doi.org/10.1017/S1355771899002046 Published online by Cambridge University Press

Figure

articles, one on information theory and the other on an ancient method of building recorders. Here is an excerpt from the third text movement:

> Extending to the notaneound-breath of the Kynch, Dancks are many come to be known from the pinkastly come to eb accessfullows openeralizably the adject (19640) based loof one of the most string point interval of the unit can sense that funced into a normal code appears of recent ty for hand-brand-baroquan adjuring of recent to the eferred insipian and Jorment out of randolfpaper [I an infless conde caneusly, Ky cussic teners. Pape end-blown the continue to produced a lourly palced with Jorma Rissanen lattered recovery of the sourcedical cirst known band the ske in which eful in 1925].

It is interesting to compare this text to the two text examples (Shannoniser and Dialectiser) presented earlier. It is obvious that our text is more Gibberish, since it creates new words not by 'copying' complete words but rather continuing existing words by segments from other words. Thus, it is also different from the Dialectiser which mostly substitutes letter combinations in existing words but does not 'invent' new words.

7. CONCLUDING REMARKS

The method of random recombination seems to provide an interesting test bed for our understanding of randomness and its relation to music schemata (Cohen and Dubnov 1997). Musically, it seems to produce better results than Markov models. This is not surprising since we do not have to limit ourselves to a certain order but choose optimally the longest possible continuation. Significant work has been done since the inception of this idea in order to provide a more compact and graspable representation of music sequences and patterns (Dubnov, Assayag and El-Yaniv 1998, Assayag, Dubnov and Delerue 1999). A new representation was recently developed that does not require complete reference pieces to be stored for the purpose of new music generation. This will be described in a future paper.

Although the beauty of the method lies in the combination of an interesting musical result with a formal theoretic backing, the research is still at a stage where more questions are posed than answered. One of the common questions that we encountered in presenting this work was: Is this a method of composition or just an intellectual experiment?

Up until now we have not fully defined or developed the methods for controlling the course of the composition (although ideas like tendency masks (Xenakis 1971) or profiles (Ames 1989) could be straightforwardly applied). The possibility of controlling expectation and surprise is obvious: one could select the least probable continuation, which is equivalent to preferring the 'uncertain' or choosing the 'surprising'. Since *NTrope Suite* was the first experiment in this direction, we did not attempt to exert our control over the process. Rather we preferred to passively observe how the music unfolds through its statistics. This process is not invalid. Contemplation can be found in the words of John Cage: 'This play, however, is an affirmation of life – not an attempt to bring order out of chaos nor to suggest improvements in creation, but simply a way of waking up to the very life we're living, which is so excellent once one gets one's mind and one's desires out of its way and lets it act of its own accord.'

An interesting observation was suggested, relating this method to contemporary ideas of deconstructivism and eclecticism. Here we have presented a tool that can create something new, but still relying on the whole past musical heritage. This means that there is some way to 'revive' the past and incorporate it in new original works, not merely as a 'quotation' but in an abstract and original sense. Probably, once the 'affirmation of life' has been tamed, composers will find it more useful. Some of them already find it useful: Jean-Remy Guedon from Orchestre National de

https://doi.org/10.1017/S1355771899002046 Published online by Cambridge University Press

Jazz wrote a piece using the computer to help write his improvisations. The program notes said: 'le sujet poetique est de fondre improvisation et ecriture grace a un algorithme mis au point par Shlomo Dubnov et Gerard Assayag dans OpenMusic, developpe a l'Ircam.' (Portes Ouvertes, Ircam, 1999).

It is probably best to conclude this paper with another quotation from Shannon's fundamental paper:

> The concept of information developed in this theory at first seems disappointing and bizarre – disappointing because it has nothing to do with meaning, and bizarre because it deals not with a single message but rather with the statistical character of a whole ensemble of messages, bizarre also because in these statistical terms the two words information and uncertainty find themselves to be partners. (Shannon 1949)

Rereading Shannon's text, substituting 'composition' for 'message' and 'meaning' for 'music', is the mix we suggest.

REFERENCES

Ames, C. 1989. Stylistic automata in *Gradient*. In C. Roads (ed.) *The Music Machine*. Cambridge, MA: MIT Press.

Assayag, G., Dubnov, S., and Delerue, O. 1999. Guessing the composer's mind: applying universal prediction to musical style. *Proc. of the Int. Computer Music Conf.* Beijing: ICMA.

Cohen, D., and Dubnov, S. 1997. Gestalt phenomena in musical texture. In Marc Leman (ed.) *Music, Gestalt and Computing*, pp. 317–36, Series of lecture notes in artificial intelligence. Springer Verlag.

Cope, D. 1991. *Computers and Musical Style*. AR-Editions.

Dubnov, S., Assayag, G., and El-Yaniv, R. 1998. Universal classification applied to musical sequences. *Proc. of the Int. Computer Music Conf.* Ann Arbor: ICMA.

El-Yaniv, R., Fine, S., and Tishby, N. 1997. Agnostic classification of Markovian sequences. In M. Jordan, M. Kearns and A. Solla (eds.) *Advances in Neural Information Processing Systems*. Vol. 1, pp. 965–71.

Hiller, L., and Issacson, L. 1959. *Experimental Music*. McGraw-Hill.

Shannon, C. 1949. *The Mathematical Theory of Communication*. University of Illinois Press.

Xenakis, I. 1971. *Formalized Music*. Indiana University Press.