HCCE '12 Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments, Pages 194-199 ACM New York, NY, USA ©2012

Prediction of Dissimilarity Judgments between Tonal Sequences using Information Theory

Michael Frishkopf
University of Alberta
382 FAB, Department of Music
Edmonton, Alberta, Canada T6G2C9
1-780-492-0225
michaelf@ualberta.ca

ABSTRACT
Several studies have investigated the entropy characteristics of musical notations (in the information-theoretic sense). What appears to be lacking is an empirical study of the connection between entropy and musical perception. This paper describes results of an experiment designed to determine the relevance of entropy to subjects' dissimilarity judgments on pairs of melodic sequences. The hypothesis was that dissimilarity judgments by subjects on pairs of unfamiliar tonal sequences drawn from a common pitch set are largely a function of average sequence entropy and average sequential interval size, when the sequences are uniform in all respects except for the ordering of pitches. Five stationary ergodic Markov-1 chains of increasing entropy were defined on a common pitch set. From each chain, two sequences of identical entropy were generated: the first sampled directly from the chain, and the second by applying to each sequence element of the first a random permutation of the pitch set. In this fashion, entropy and average interval size variables could be varied quasi-independently. Timbre, duration, and loudness were held constant. Subjects heard all possible unordered pairs of synthesizer-generated sequences through headphones, and indicated a subjective dissimilarity rating for each pair. Two forms of analysis yielded different results. Subject dissimilarity judgments between sequences were shown to be well correlated with a Euclidean distance function on average interval size and entropy. However, multidimensional scaling analysis revealed only average interval size to be a salient judgment factor, not entropy.

Categories and Subject Descriptors
J.4 [PSYCHOLOGY]; J.5 [ARTS AND HUMANITIES]: Music.

General Terms
Algorithms, Measurement, Experimentation, Human Factors.

Keywords
Information, entropy, perception, cognition, music, melody.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
*HCCE'*12, March 8–13, 2012, Aizu-Wakamatsu, Fukushima, Japan.
Copyright 2012 ACM 978-1-4503-1191-5 …$10.00.

1. INTRODUCTION
The experiment described in this paper is intended as a small step towards understanding the relevance of information-theoretic entropy in music perception.

Shannon and Weaver [7] first quantified information as entropy within a broader mathematical theory of communication, with the practical aim of improving the quality of telephone circuits.

But their elegant theory soon found applications beyond electrical engineering, becoming especially fashionable (in conjunction with the burgeoning field of semiotics) in the social sciences and humanities. Shannon himself examined the information content of the English language [8], while humanists including Gregory Bateson [1] and Leonard Meyer [4,5] attempted to apply information-theoretic reasoning to the arts. But these arts applications were largely speculative, lacking empirical support.

The earliest empirical information-theoretic studies in music seem to have been carried out by Youngblood, who analyzed songs by Schubert, Schumann, and Mendelssohn, aiming to show the relation between musical style and entropy [8]. Subsequently, Knopoff and Hutchinson published papers suggesting that continuous musical variables be examined for entropy, using entropy analysis to differentiate musical styles [2,3]. However, these empirical studies center on notation: the entropy of the symbols in musical scores are correlated with generally accepted categories of musical style and aesthetic excellence. What appears to be lacking is scientific research on the relation between entropy and musical perception. The present study is a step in this direction.

2. HYPOTHESES
The entropy of a symbol sequence depends upon both the sequence, and the underlying stochastic process by which it was generated. Given a stochastic process alone, one can also calculate the expected entropy per symbol, as a weighted average over all possible sequences. But to speak of the entropy of a symbol sequence alone is meaningless. Among the equiprobable set of all grammatical English word sequences of equal length, the Bible contains tremendous information. However, considered as one of a small number of scriptures, the entire Bible can be represented in a handful of bits.

Thus, entropy applied to musical sequences is only meaningful if we can assume some probabilistic model of musical expectation in the listener's mind. To exemplify with a simple instance: imagine a game of musical chairs whose music consists of a monophonic sequence of equidurational tones. When this music is suddenly paused, what probabilities does the listener assign to the subsequent tone? Less predictability implies greater entropy.

This assumption of a probabilistic mental model, adopted by Meyer and others, is fairly uncontroversial, and is widely understood to underlie, at least in part, absolute music's aesthetic power. But how might such a model be formed? In general, it would appear to depend on the individual's entire life history of musical exposure, during which she or he has been conditioned to particular musical styles and genres. Such exposure must cause the accretion of a relatively stable, slowly changing, musical model, used to formulate musical expectations across a variety of listening contexts.

However, within any particular listening context, we hypothesize that the listener dynamically augments this stable model with a more ephemeral model, based on the actual frequencies of currently perceived musical events, which is revised (or reinforced) as it is contradicted (or confirmed) by the auditory stream. The more the perceived music is unfamiliar, and hence beyond the scope of the stable model, the greater the importance of the dynamic model. Such, at any rate, is our hypothesis.

At the admitted risk of oversimplification, we may represent this situation mathematically as follows: We suppose the listener to parse the auditory stream into a linear succession of symbols.[1] The listener continually averages certain musical statistics within a window constrained by short-term memory, such as the frequencies of pitches, pitch transitions, contours, and so forth. The shifting collection of statistics is the dynamic model, and thus provides one basis for entropy. With unfamiliar music, it is the primary basis.

It is not our intention to establish these hypotheses empirically in this experiment: we will not determine the size of the averaging window, or the statistics computed, via experiment. In any case, musical experience is clearly nothing so simple. Rather, these statements model a black box (the auditory system), serving to explicate the genesis of the experimental hypothesis, to motivate experimental design and analysis, and to suggest future work.

The hypothesis that we put to empirical test is the following:

Dissimilarity judgments by subjects on pairs of unfamiliar tonal sequences drawn from a common pitch set are largely a function of average sequence entropy and average sequential interval size, when the sequences are uniform in all respects except for the ordering of pitches.

Specifically, we propose that the dissimilarity between two such sequences, α and β, is Euclidean, according to the predictor:

$$d(\alpha,\beta) \approx \sqrt{(c\Delta AI)^2 + \Delta E^2}$$

where:

$d(\alpha,\beta)$ = dissimilarity judgment between α and β
ΔAI = difference in average interval size between α and β
ΔE = difference in average entropy between α and β
c = constant to be determined

Here, sequence entropy is defined relative to a probability model derived from sequence statistics. The problem for this experiment was to test this hypothesis, and determine a suitable constant, c.

[1]This assumption is evidently incorrect; perceived musical structure is far more complex, involving multiple tonal, harmonic, and rhythmic structures, existing in hierarchical and non-hierarchical relations to each other, but the assumption will prove useful for the synthetic, monophonic musical sequences used in the experiment.

3. METHODS

3.1 Generation of Markov Chains

Tonal sequences were derived from Markov-1 chains,[2] defined on a common set of 14 symbols representing the pitch set, PS, as shown in Figure 1:

Figure 1. Pitch set (PS) = {F3, G#3, A3, Bb3, C4, D4, Eb4, F4, Gb4, A4, Bb4, C5, Db5, E5}

This pitch set was chosen for two reasons: (a) it does not replicate at the octave, and (b) it contains intervals that are unusual by Western standards. PS is thus likely to sound unfamiliar for experimental subjects drawn from the student population at UCLA, where the experiment was performed.[3] Therefore, use of PS promotes relatively unfamiliar musical sequences, a situation which is desirable according to our hypotheses, because such sequences should place greater emphasis upon the dynamic musical model constructed during the listening experience, reducing emphasis on the more stable long-term model resulting from a lifetime of listening, and increasing the likelihood that dissimilarity judgments may be predicted based on attributes of auditory stimuli alone.

A Markov-1 chain can be defined by giving an initial probability distribution P_0, and a constant transitional distribution M, where $P_0(i)$ is the probability of starting the sequence with symbol i, and M(i,j) is the conditional probability that j will be the subsequent symbol, given that i is the current symbol.[4] Letting P_n be the probability distribution at time slice n (n=1,2,3…), it is easy to prove the Chapman-Kolmogoroff equation: $P_n = P_0M^n$ (where multiplication and exponentiation are matrix operations), or (equivalently) $P_n = P_{n-1}M$.

Two critical properties of Markov chains are stationarity, and ergodicity. If $P_n = P_m$ for all n and m, then the chain is stationary: its statistics are not changing over time. Given M, it is possible to choose P_0 such that the chain is stationary. Therefore, assuming stationarity, the chain is defined by giving M alone, since $P = P_0$ can be derived from M (indeed, as P = PM, it is clear that P is simply a left eigenvector of M with unit eigenvalue). The chain is called ergodic when the statistics of any sample sequence reproduce the statistics of the chain itself. To ensure ergodicity, it is necessary and sufficient that for any pair (i,j), $M^n(i,j)$ is positive for some n: any element can be reached from any other with non-zero probability, after some number of transitions.

All chains used to generate the sequences for this experiment were constructed to be both stationary and ergodic. These requirements follow from the underlying hypotheses. If the chain is nonstationary, then the mental model resulting from averaging the heard sequence over a time window would be changing. Since we have no idea how long such a window might be, and since we don't want the subject to become confused by changing averages, it is preferable to ensure that statistics be constant. If the chain were non-ergodic, then a particular symbol sequence—however

[2]In a Markov-1 chain, the probability of the chain assuming a particular value at integral time n depends only upon the value assumed at time n-1.

[3]This scale is derived from *saba*, a mode (*maqam*) used in Arab and Turkish music. See http://www.maqamworld.com.

[4]Thus the rows of M always sum to unity.

long—would not represent the chain from which it was drawn, statistically speaking; it might possess a totally different entropy. Since we want to control sequence entropy by controlling chain entropy, non-ergodicity is not permissible.

Markov chain entropy is defined as usual:

$$H = \sum_i \sum_j P(i) M(i,j) \log_2 \frac{1}{M(i,j)}$$

As the distribution of symbols tends towards equiprobability, entropy increases. Therefore, there is a simple smoothing technique for increasing the entropy of a chain defined by transitional distribution matrix M: one adds a constant to every element of M, and renormalizes by row. Clearly, when such a constant is large, elements of M all become equal. In this case, the corresponding P required to ensure stationarity will also consist of equal elements, and entropy will be maximized.

Thus, in order to generate the Markov-1 chains for this experiment, the following procedure was used:

- A sequence in pitch set PS was composed. This sequence deployed principally conjunct melodic movement.
- Transition statistics for the composed sequence were computed, creating a matrix M.
- Chain M_0 was constructed, based on M.
- Chains M_1 – M_4 were constructed, by smoothing M_0 using the technique described above.

The result was the set of five chains shown in Table 1, where entropy denotes average bits per symbol, and entropy differences are computed relative to the previous chain.[5]

Table 1. Markov-1 chains and their entropies

Chain	Entropy (bits)	Entropy difference $M_i – M_{i-1}$
M_0	1.512342	-
M_1	2.087372	0.57503
M_2	2.662557	0.575185
M_3	3.431763	0.769206
M_4	3.807145	0.375382

3.2 Generation of Tonal Sequences

A sample sequence, $S_k(n)$ of 225 pitches was extracted from each of the five chains, M_k (k=0,...,4; n=0,...224). Next, five different permutations π_k of the 14 pitches were randomly generated, and applied to the five sequences respectively to create five additional, permuted sequences, as follows:

$$S_{k+5}(n) = \pi_k (S_k(n))$$

Entropy, and average interval size (average one-step transition size, in semitones), were computed for each of the 10 sequences, based on a matrix M and distribution P derived by counting the frequencies of pitches and pitch transitions for each sequence. Note that the permutation operations do not affect entropy, i.e.:

$$\text{entropy}(S_{k+5}) = \text{entropy}(\pi_k (S_k)) = \text{entropy}(S_k)$$

Average entropy and average interval size are not independent variables. When entropy is high, average interval size cannot be low, since all intervals must be used equiprobably. However, when entropy is low, average interval size may be high or low. Permuting the pitches in a sequence with low average interval size tends to increase its average interval size. Thus, permutations allow one to partially disentangle entropy from average interval size. Without such controls, it would be impossible to differentiate the effects of each variable in the experiment.

Thus, a set of 10 sequences was obtained, as shown in Table 2, where each of the five original sequences is followed by its permutation containing identical entropy but higher average interval size. Again, entropy is measured in average bits per symbol; average sequential interval size (Avg Int) is measured in semitones.

Table 2. Sequences, entropies, and average interval sizes

Sequence	Entropy	Avg Int
S_0	1.421024	1.756322
S_5	1.421024	8.736595
S_1	1.815737	2.517970
S_6	1.815737	6.846851
S_2	2.201215	3.652956
S_7	2.201215	8.279984
S_3	2.759830	5.096834
S_8	2.759830	7.134021
S_4	3.179414	7.023991
S_9	3.179414	7.809772

Musical variables other than pitch were kept constant. All sequences were played by a computer-controlled synthesizer, at a constant rate of 80ms/note, at fixed loudness[6] and timbre.[7]

3.3 Experimental Procedure

Seven subjects were employed, including one who participated twice, for a total of eight experimental sessions. All subjects had some musical training; all except the double participant were music majors at UCLA. Subjects heard the sequences, generated by a MIDI-controlled synthesizer, through headphones. Subjects were instructed to enter an integer from 0 to 99 for each pair of sequences presented, where 0 indicates minimal dissimilarity (identity), and 99 indicates maximal dissimilarity. They were urged to listen to the entire sequence, to use the entire scale 0-99, and to attempt to formulate a judgment strategy and to apply it consistently.

First, all ten sequences were presented, in random order, so that subjects could get a sense of their diversity. Next, from four to eight practice pairs were presented, drawn at random from the set of all unordered pairs.[8] During this practice period, subjects

[5]Note that the maximum possible entropy on 14 symbols is $\log_2 14 \approx 3.81$. Also note that entropy increments are unequal. There is a subtle relation between the entropy of a chain, and the entropy of a sequence drawn from the chain, defined in terms of its own statistics. It was necessary to adjust the chain entropies so as to arrive at a near equal spacing of sequence entropies.

[6] Loudness was fixed except for a brief (1.35 s) ramp up from, and down to, zero at the beginning and ending of each sequence. Ramping of loudness was intended to eliminate any extra significance attached to the first or last pitch.

[7]The rapid tempo necessitated a sharp attack, with decay to the next note. A marimba-like sound was found to be suitable.

[8]The number of practice pairs was increased when it was realized that they were being used effectively to formulate a judgment strategy.

learned how to operate the equipment, and began to formulate a judgment strategy. Finally, all 55 unordered pairs (including identities) drawn from the ten sequences were presented once each, in random order. Subjects were given a brief rest every ten pairs. A single run of the experiment required between 40 and 50 minutes. At the conclusion of each run, the subject was queried as to the strategies he or she had used in making dissimilarity judgments.

4. ANALYSIS AND RESULTS

4.1 Statistical Analysis

The dissimilarity judgments produced by each subject during the course of a single experiment run form a lower-triangular matrix, what we will call a "triangle," whose elements all fall into the range $\{0,\ldots,99\}$. Eight triangles, $t_0 - t_7$, each resulting from a single experiment run, were averaged to produce a ninth triangle, t_8. In addition, two completely random triangles (t_9 and t_{10}) were constructed.

An algorithm was designed and implemented as a computer program to construct synthetic triangles, $t_s(c)$, using a Euclidean metric d_s to define the distance between melodic sequences α and β,

$$d_s(\alpha,\beta) \equiv \sqrt{(c\Delta AI)^2 + \Delta E^2}$$

where (as before) ΔAI = difference in sequence average interval size, and ΔE = difference in sequence entropy between α and β. For each t_i this program seeks an optimal value (c_i) for c that maximizes the Pearson correlation between triangles t_i and $t_s(c)$, returning the correlation as r_i. Searching 60 logarithmically equidistant values of c between 1/1000 and 1000 resulted in approximate values for c_i as given in Table 3.

Table 3. Approximate c_i and r_i for each triangle

Triangle	$Log_{10}(c_i)$	correlation r_i
t_0	1.7000000	.6359944
t_1	-0.4000006	.6013461
t_2	-0.1000006	.7988621
t_3	-0.3000006	.6740924
t_4	-0.3000006	.8131643
t_5	-0.6000006	.4820028
t_6	-0.0000006	.8559344
t_7	-0.4000006	.7491813
t_8	-0.3000006	.8954409
t_9	2.899999	-.06348647
t_{10}	-0.1000006	.03700948

Note that it is possible to attain fair correlation for all triangles except the random ones (t_9 and t_{10}), an encouraging result. In particular, triangles 2,4,6, and 8 (the average triangle) show very good correlation. These results indicate that entropy and average interval size are likely factors in subjects' dissimilarity judgments, although somewhat differently scaled for each. The optimal c value for subject 0 is strikingly different from the others. This anomaly is in accord with the fact that subject 0 had some difficulty with the test (having accidentally inverted the dissimilarity scale for a few judgments, according to her report), and was the only non-music major.

Next, the mean log value of c was used to construct a single synthetic triangle t_{11}, and correlation coefficients r_{ij} were computed between triangles i and j, for i = 0,...,11, and j <= i. The resulting triangular matrix r_{ij} was converted to a dissimilarity matrix, using the formula $d_{ij} = 20 - 10(r_{ij}+1)$,[9] which was mapped into two dimensions using classical multidimensional scaling (MDS). The resulting plot, shown in Figure 2, is hardly instructive except to indicate that all the non-random triangles are relatively close together compared to the random triangles. Clearly, subjects' performance, and the synthetic dissimilarity measure, are significantly in agreement.

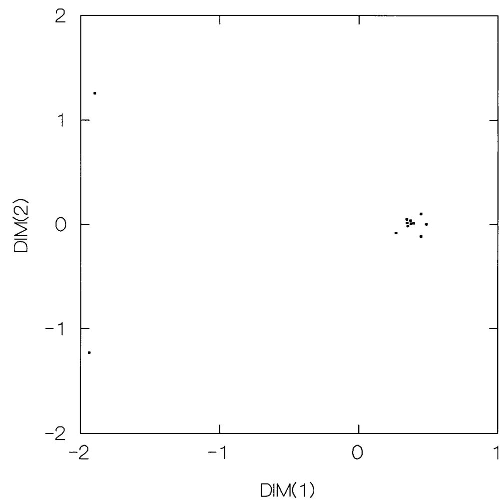

Figure 2. MDS of subject triangles, average triangle, synthetic triangle, and random triangles (outliers). (Stress = 0.013, proportion of variance = 100%)

In the next stage of analysis, the random triangles were discarded, and a new average triangle t_8 was constructed from triangles t_1 to t_7, since the data in t_0 was known to be contaminated. The synthetic triangle, $t_9 = t_s(0.5)$ was chosen so as to maximize correlation with the new average triangle t_8. (Note that 0.5 corresponds to -0.3 on the log scale used above.) Table 4 shows resulting correlation coefficients r'_i between triangles t_i and $t_s(0.5)$ (the third column indicates the correlation degradation suffered by using c = 0.5 rather than the approximately optimal c_i value for each triangle, as given in Table 3):

Table 4. Correlation coefficients r'_i between t_i and $t_s(0.5)$, and correlation degradations $r_i - r'_i$

Triangle	r'_i	$r_i - r'_i$
t_0	0.5999775	-3.601694E-02
t_1	0.5962809	-5.065262E-03
t_2	0.7923717	-6.49035E-03
t_3	0.6740925	5.960464E-08
t_4	0.8131641	-2.384186E-07
t_5	0.4499896	-3.201315E-02
t_6	0.8367776	-1.915675E-02
t_7	0.7460104	-3.170907E-03
t_8	0.8709132	-5.960464E-08

Since the deviations from optimal are low, we concluded that c=0.5 is a satisfactory value for the entire subject population.

[9]This formula maps a correlation of +1 to 0, and -1 to 20; thus high correlation becomes low dissimilarity, and vice versa [6].

As before, correlation coefficients were computed for each unordered pair of triangles, including the synthetic triangle t_9, and the results converted to dissimilarity data and scaled in two dimensions using MDS. The ten triangles (t_0 - t_9) were also subjected to Systat's cluster analysis, using complete linkage, and using correlation coefficients as the distance metric. The results are shown in Figures 3 and 4 (note that triangles t_0 - t_9 are here relabeled a-j, respectively). Clustering provided by Systat are also indicated on the MDS plot.[10]

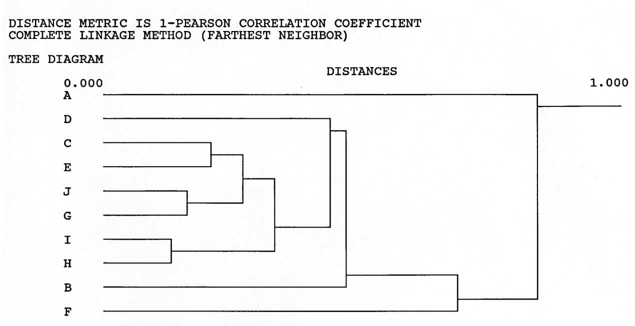

Figure 3. Cluster analysis, using complete linkage and Pearson correlation distance metric, of subject triangles (a-h), average triangle i (excluding t_0), and synthetic triangle j.

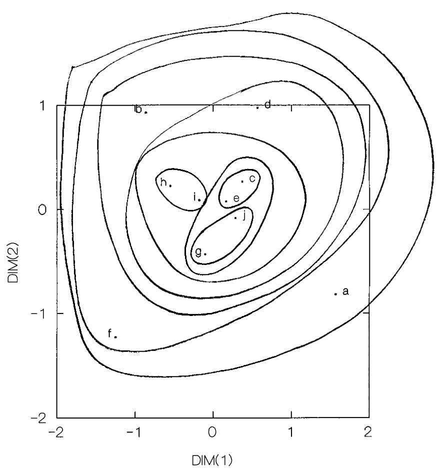

Figure 4. MDS of subject triangles (a-h), average triangle i (excluding t_0), and synthetic triangle j. Cluster analysis is superimposed. (Stress = 0.087, proportion of variance = 97%)

Understandably, the synthetic triangle is not an outlier points, but rather is deeply nested in one of the two innermost clusters. The clustering also shows that t_0 (a) is an outlier, as was already suspected.

Finally, individual MDS plots were constructed for all the triangles, mapping melodic sequences (S_0 – S_9) in two dimensions. With suitable rotations of these plots, especially for t_1, t_4, t_6, t_7, the projection of each point onto dim(1) is clearly correlated with average interval size (see Figure 5; note that melodic sequences S_0 – S_9 are here relabeled a-j, respectively). However, the effects of entropy are more difficult to see in the plots, except for the synthetic triangle. This result is puzzling, since the result c=0.5 seems to indicate an important role for entropy in dissimilarity judgment. It appears that other factors may be more salient than entropy in formulating dissimilarity judgments, and further investigation is indicated.

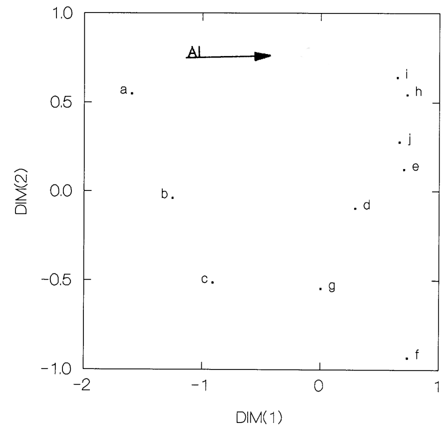

Figure 5. MDS of melodic sequences using the average triangle (excluding t_0), indicating direction of increasing average interval size (AI). Direction for increasing entropy is unclear. (Stress = 0.14, proportion of variance = 90%)

4.2 Subject Strategies

Following each session, subjects were queried for judgment strategies, as a technique supplementing etic statistical analyses with emic qualitative research. The results are valuable for suggesting refinements to the predictor, and in designing future experiments.

These brief interviews indicated that subjects tend to categorize the sequences, and to judge dissimilarity partly by category (several subjects indicated dissatisfaction with the 'continuous' scale of 100 steps, suggesting that 10 or fewer would be sufficient). Categories were based upon interval size (conjunct versus disjunct motion), fissioning (whether or not the melody split into multiple voices), musical recognition (whether or not the melody contained some memorable musical characteristic), range (high or low), and scale (although the pitch set was always the same, different chains stress different notes, and thus lead to the perception of different ranges or scales). All subjects reported difficulty distinguishing identities;[11] this task was easiest when sequences presented some memorable melodic motif.

The existence of melodic fissioning, mentioned by most of the subjects, was not important merely as a factor in categorization, for subjects also reported listening to fissioned melodic lines within a perceptual polyphonic texture, and evaluating them independently. For instance, two subjects reported focusing primarily on an upper line. Thus, fissioning, by partitioning the sequence into two or more concurrent sequences, changes the statistics of note and transition frequency in a way that the predictor did not take into account.

[10]In these figures, the triangles are relabeled a-j.

[11]This difficulty was reflected in their triangles (maximum identification was 80%).

5. CONCLUSIONS

In this experimental setting, of admittedly limited musical validity, dissimilarity judgments between sequences are well correlated with a Euclidean dissimilarity predictor:

$$d(\alpha,\beta) \approx \sqrt{(0.5 \times \Delta AI)^2 + \Delta E^2}$$

where:

$d(\alpha,\beta)$= dissimilarity judgment between α and β
ΔAI = difference in average interval size between α and β
ΔE = difference in entropy between α and β

However, MDS analysis shows that average interval size is a judgment factor, while the role of entropy is unclear.

Interviewing subjects on their judgment strategies revealed that average interval size was an explicit factor, that judgments may have been more categorical than continuous, and that melodic fissioning was an important perceptual phenomenon impacting their dissimilarity judgments.

6. FURTHER RESEARCH

Many questions arise from this small study. First, it would be valuable to investigate whether judgments are made consistently, and whether they are really symmetric ($d(\alpha,\beta) = d(\alpha,\beta)$), as was assumed here. To answer these questions would require more extensive experimentation, in which subjects judged all ordered pairs of stimuli, with replication. Any asymmetry would present an interesting problem for study. The experiment should also be replicated to determine the effects of sequence length, scale, and tempo.

One could also perform a similar experiment, in which sequences were generated randomly anew from the Markov chain at each presentation; the question would then be: can dissimilarity judgments be predicted from the chain itself? Such an experiment might diminish the importance of melodic recognition, which is a distraction from the central issue here.

Related experiments of interest include the following: (1) Semantic differentials. Which differentials, if any, correspond to entropy? “emotion,” “complexity,” “randomness”? (2) Perception of entropy. Can subjects be trained to estimate sequence entropy? What is the just-noticeable-difference curve for entropy? (3) Production of entropy. Can subjects be trained to produce sequences of entropy E, given sufficient exposure to sequences of the same entropy level? (4) Higher order chains. Can use of higher-order Markov chains lead to an improved dissimilarity predictor? (5) Categorization. When subjects categorize melodic sequences according to their own criteria, do the resulting categorizations bear any relation to entropy?

In addition, subjects’ reports indicate some room for improvement in the synthetic dissimilarity measure. The importance of melodic fissioning implies that the entropy measure ought to account not only for adjacent note transitions, but also transitions between notes in the same register. Such a quantity should be calculable from the sequences, or even from the Markov chains themselves.

7. ACKNOWLEDGMENTS

Our thanks to the anonymous subjects for their patient dedication, to the UCLA Systematic Musicology Laboratory for providing experimental facilities, and to Professor Roger Kendall for his guidance.

8. REFERENCES

[1] Bateson, Gregory. 1972. Style, Grace, and Information in Primitive Art. In *Steps to an Ecology of Mind*. Chandler Pub. Co., San Francisco.

[2] Knopoff, L. and Hutchinson, W. 1981. Information Theory for Musical Continua. *Journal of Music Theory* 25, 1, 17-44.

[3] Knopoff, L. and Hutchinson, W. 1983. Entropy as a Measure of Style: The Influence of Sample Length. *Journal of Music Theory* 27, 1, 75-97.

[4] Meyer, L. B. 1956. *Emotion and Meaning in Music.* University of Chicago Press, Chicago.

[5] Meyer, L. B. 1957. Meaning in Music and Information Theory. *The Journal of Aesthetics and Art Criticism* 15, 4.

[6] Senju, M. and Ohgushi, K. 1987. How Are the Player's Ideas Conveyed to the Audience? *Music Perception*, 4, 4, 311-323.

[7] Shannon, C.E., and Weaver, W. 1949. *The Mathematical Theory of Communication.* University of Illinois Press, Urbana.

[8] Shannon, C.E. 1951. Prediction and Entropy of Printed English. *Bell System Technical Journal* 30, 50-64.

[9] Youngblood, J.E. 1958. Style as Information. *Journal of Music Theory* 2, 24-35.