Cohen, Joel E., ***Information Theory and Music***, Behavioral Science, 7:2 (1962:Apr.) p.137

Jonathan Swift once said that good writing was a matter of getting the right words in the right order. Similarly, a composer's task can be described as selecting among available sounds and arranging these sounds into some sort of pattern. The uses of information theory to analyze and describe this creative process, its limitations and possibilities, are discussed in this paper.

INFORMATION THEORY AND MUSIC[1]

by Joel E. Cohen

Harvard University

A SYNTACTIC STUDY OF MUSIC

THE classical presentations of information theory were technical. The theory was developed for communication media where certain assumptions about the signs transmitted through the media could be made, and where the communication problem was defined as the effective transmission of physical sign-vehicles. Originally the domain of the theory as a scientific theory was properly restricted.

But the fundamental concepts of the theory have a great intuitive appeal. In the last decade, information theory has been applied in at least a dozen different fields. In some extensions, the use of the calculus of information theory was carefully justified or the calculus was modified according to the requirements of the field of study. In other extensions, however, "experiments" were performed without regard to their validity or significance. This was usually done by appealing to the reader's intuition with amorphous generalities, then leap-frogging to the H-formula (Shannon, 1957) for information-content and inserting some numbers. Of this trick, extensions into musical theory have been particularly guilty.

Musicians with a taste for mathematics have applied the H-formula without regard to the assumptions they were making. Technicians have applied informational measurements without regard to aesthetic considerations. To the degree that these interfield minglings have stimulated objectivity in examining musical "intangibles," they have been helpful; but when validity is claimed for their results, inquiry should be made into the means of obtaining them.

This paper is an attempt (1) to explicate some assumptions already made in previous applications of information theory to music; (2) to survey the results obtained on the basis of these assumptions; and (3) to criticize these results and assumptions in order to indicate some necessary modifications and possible new areas of research.

Syntactic theory

Tischler (1956) has shown that the relations which give a work of art aesthetic value may be divided into two classes. The first class consists of syntactical, or internal, relations, which are unique to the medium of the aesthetic object. In music, they include rhythm, melody, harmony, counterpoint, tone color or instrumentation, expression or interpretation, and form or "meaningful contour." The second class consists of nonsyntactical, or external, relations, which affect all arts. They include gesture, programme, "ethics or the expression of the composer's emotions and ideas, and his technical mastery," the drives and motivations of the creator as reflected in his work, intended social functions of the work, historical and social conditions, the place of the work in the history and tradition of the art, the personality and proficiency of the performers, if any, and the conditions of performance, if any.

The ultimate goal of a syntactical study of music is to understand how music communicates. The value of this limitation to

[1] This paper was written at Cranbrook School, Bloomfield Hills, Michigan.

the *how* of communication is that the results of syntactical study do not depend on the many speculations about *what* music communicates, and may be valid whether or not music communicates emotions, "significance" (Langer, 1952), emotional images, "divine harmony," moral states, or nothing at all beyond itself.

To place this syntactical study of music in proper relation to other studies of communication, semiotic, the theory of signs, provides a useful framework and metalanguage (Morris, 1955).

Musical sign systems

A basic problem in the syntactical study of music is the determination of "sign-vehicles." There is a sequence of perceptions in the mind of a listener, measured inferentially in psychology. There is a sequence of events in the air or transmission cable, measured in physics. There is an operational schema, the "score" of a "piece of music," representing certain aspects of the psychological and physical events. Each of these sequences forms an interconnected system of signs. Each sign system is closely related to the others.

So far most scientific study in music has been devoted to the alphabets of the psychological and physical sign systems and the relationships between them. Psychologists have determined the upper and lower perceptible thresholds and the least perceptible differences of frequency, amplitude, and time. Physicists have studied the structure of complex sounds, analyzed the timbres of instruments in varying combinations, and increased the efficiency of sound transmission. Helmholtz (1954) established that the physical correlate of the psychological phenomenon timbre (the distinctive "color" of the oboe or violin tone, for example) is the spectrum or wave form of the sound. The Weber-Fechner law very approximately relates pitch (psychological) and frequency (physical), loudness (psychological) and amplitude (physical), by the general law, $S = K \log E$, where S is the psychological sensation, E is the physical excitation, and K is a constant (Moles, 1958).

The psychological sign system of music gives a little less information (both in the general sense and in the technical sense) about a given performance of music than does the physical description. Whereas the measuring instruments of physics can determine the frequency of a sound to a small fraction of a cycle per second, the human receptor can distinguish only about 1,200 distinct pitch levels within the audible range of 16 to 16,000 cycles per second. He is insensitive to alterations below a differential threshold on the order of one comma. Similarly insensitive below a differential threshold in amplitude on the order of one decibel, the human receptor distinguishes about one hundred levels of loudness, other things (pitch, duration, etc.) being equal. The physically distinguishable sounds which are psychologically the same are called "metameric" (Meyer-Eppler, 1958, p. 59). Since there are apparently many metameric sounds, the basic alphabet of signs from which the physical sign system draws is much larger than the alphabet from which the psychological sign system draws.

An acculturated receptor of a musical performance tends to ignore differences which he is physiologically able to distinguish, in order to recognize the structure of the composition, although he may use all his available discrimination in order to evaluate a performance. For example, a listener may be able to recognize Beethoven's *Fifth Symphony* even when transcribed for two pianos and poorly performed. He may recognize that a singer intended to sing a certain pitch even though she is as much as a quarter tone off. Listeners can follow a metric pulse in time through extended *rubatos* and *accelerandos*. The pattern that is usually thought of as a piece of music is not, then, a single sequence of sounds or of sensations, but is rather a "field of liberty," a set of many possible physical and psychological realizations.

The score of a piece of music represents the "field of liberty." The elements of the score (pitch, rhythmic indications, etc.) are the "cultural alphabet," general directives with which composers have had to content themselves in the past. The cultural alphabet is the set of culturally significant sounds available to the composer. It is analogous to the phonemic reduction of a spoken language, a set of symbols which represent the culturally significant speech sounds in a language community (Moles, 1958).

Flexible mappings relate the physical, psychological, and cultural sign systems of music. These dynamic mappings change constantly in the listening process. In the mapping between the physical and psychological systems, the ear subsumes physically distinguishable but metameric sounds together under the rubric of one distinguishable pitch (or loudness or duration) level. In the mapping between the psychological and cultural systems, the acculturated listener lumps together distinguishable pitch, loudness, and duration levels into one culturally significant sound unit, represented in the score by a written note.

Mathematically, since a many-one mapping applied to the psychological sign system gives a structure isomorphic to the cultural sign system, the cultural sign system (the score) is a homomorphism of the psychological sign system.

This formulation of the relation between the psychological and cultural sign systems seems to resolve the pattern-performance dichotomy of Springer (1956), who observes that music seems to have both a concrete physical level existing in the air between the producer and the perceptor, and an "idealized, abstract level" existing in the minds of both. According to the formulation given, the idealized, abstract level is simply the cultural sign system created when the producer and perceptor properly ignore enough details in the psychological sequence or message.

Applications of information theory to the study of the syntactic structure of the cultural sign system will be reviewed and evaluated here.

NECESSARY RESULTS FROM INFORMATION THEORY

Many applications to music which present the intuitive approach to Shannon's measure of information-content never bother with the formal approach. This gives the unfortunate impression that since the theory itself appears vague and cloudy the results of its application must have little validity.

The several mathematical assumptions about the nature of the sources to which information theory can be applied will be reviewed below in the critique of the applications. One assumption may be noted: since all applications so far have considered compositions as sequences of discrete cultural sign-vehicles, they have used only the theory of discrete sources with finite alphabets. The general theory also deals with the information-content of continuous wave forms; it may be that future applications will consider the wave forms of music as letters or sign-vehicles.

An adequate formal development of the discrete theory (Khinchin, 1957) defines Shannon's H-formula in terms of a finite scheme, a set of mutually exclusive possibilities with associated probabilities.

The properties of the H-measure which are relevant to musical applications may be briefly reviewed.

If all the probabilities in a finite scheme are zero except one, p_k, and $p_k = 1$, then there is no uncertainty, since the associated event E_k is bound to occur; appropriately $H_1 = 0$.

If all the probabilities in a scheme of n events are equal, uncertainty is maximal; appropriately H then takes its greatest value ($H_1 = \log n$).

If a finite scheme B is dependent on A, so that the occurrence of A_k changes the probabilities in B, then the occurrence of A_k can only reduce the uncertainty associated with B. As a result, in a sequence of events whose schemes are dependent on preceding events, the additional uncertainty of each new event tends to diminish. (Cf. p. 160.)

Shannon's measure of information-content is in no way a measure of "meaning." Whereas H_1,

$$H_1 = \sum_{i=1}^{n} p_i \log p_i, \qquad (1)$$

applies to ensemble of events or messages, "meaning" applies to a single message (MacKay, 1955).

In fact, in this new theory the word information relates not so much to what you *do* say, as to what you *could* say. That is, information [-content] is a measure of your freedom of choice when you select a message (Weaver, 1955, p. 100).

As such, while it is certainly inadequate to deal with the semantic dimension of a sign system, it is appropriate for studying the syntactic dimension.

Stochastic sources

In information theory, the output of any information source such as communicant A is considered as a stochastic process, i.e., a random source emitting signs according to probabilities. The statistical structure of the source is its mathematical definition.

Markov chains are stochastic sources in which sequential dependencies exist among the letters of the output-sequences. An mth-order Markov chain is defined by the probability of every possible (m-1)-gram (sequence of m-1 letters) and a matrix of transition probabilities from each (m-1)-gram to each letter in the source's alphabet. In an mth-order Markov chain, mth-order sequential dependencies exist; the probability of a given mth letter is influenced by the past sequence of m-1 letters. Theoretically, in a sequence of sufficient length, sequential dependencies of any order may exist.

The information gained when a Markov chain moves one letter ahead, that is, the average information-content per letter, cannot be measured by the H_1-formula alone. Considering second-order dependencies, the information gained from the second letter of a digram equals the total information-content of the digram (i,j) minus the information-content of the first letter (i). Hence the average information gain is

$$H_2 = H_i(j) = H(i,j) - H_1. \qquad (2)$$

Similarly, the average information-content of a letter when mth-order sequential dependencies are considered is noted by H_m. $H_o = \log n$ takes account only of n, the number of letters in the alphabet.

Redundancy

Corresponding to each H_m is a redundancy R_m,

$$R_m = 1 - H_m / \log n. \qquad (3)$$

This redundancy has no pejorative connotations.

To the degree that a sequence is redundant, it is in some manner regular, or lawful. Any order of redundancy above the first implies that the events are more or less patterned; that *sequential dependencies* exist among them. But neither the degree nor the order of the redundancy tells us specifically what *kind* of lawfulness, or patterning, is involved (Attneave, 1959).

Redundancy provides a quantitative measure of order.

Redundancy in the signal from transmitter to receiver serves to combat errors introduced by noise in the channel, insuring that communicant B extracts the proper message from the signal.

The concept of noise is extremely general; it includes static, distortions, disturbances, perturbations, and errors introduced into the signal. Noise changes the signal from a more probable to a less probable, and from a more certain to a less certain, state; the information-content of the message is increased. The paradox that noise should increase the information-content of a message is resolved by distinguishing between desirable and undesirable information; the problem of the receptor is to distinguish the spurious information-content due to noise from the intentional information-content from the information source. In music, "the presentation of internal relationships is held to be wholly intentional on the part of the composer, except for weaknesses" (Tischler, 1956, p. 204). Hence listeners may carefully study music of the same internal structure as noise.

Redundancy also gives "information" about the characteristics of the source. If, for instance, the final chord of a Beethoven symphony can easily be predicted, it still gives information about the structure of Beethoven as a stochastic source; there is no reason to leave off the last chord because it is redundant (Youngblood, 1958, p. 29).

As already observed, a sequence may theoretically exhibit patterning of any order. The final word of a novel depends to some degree on every preceding word. In practical applications, however, two considerations limit the order m of patterning studied. (1) As m increases, H_m may level off (approach an asymptote) or may not change at all. (2) The computational labor involved in attempting to determine probabilities of m-grams varies as n^m, where n is the number of letters in the alphabet (Attneave, 1959, p. 22). For instance, the number of possible 8-grams for a string quartet is on the order of 10^{64}, and it is likely that the patterning involved is of a much larger order than 8. Even though the wildest string quartet is probably too redundant, and certainly much too short, to have a number of 8-grams approaching 10^{64}, the computational labor is still tremendous. In some past analyses computers have played a prominent role; they will have to be used increasingly as the amount and order of analysis increases.

MUSIC AS A MARKOV CHAIN

Composition as selection

The composer has a large alphabet of possibilities. This alphabet is not a simple set of seven notes, but is the 350,000 differentiable sounds in the full range of human hearing. If the composer is writing in the Western European tradition, he restricts himself to about one hundred areas of pitch (notes), about nine relative degrees of loudness, and the many timbres and combinations of timbres of the available musical instruments. This immense restriction of his alphabet still yields a huge number of different sound possibilities for any given point in time, and this huge number increases exponentially with each successive point in time. For composers demanding an even larger alphabet, such as the electronic-music composers, the problem of selection is increased many times. In short, the composer is presented with a plethora of musical materials.

The composer's task is to select among available possibilities and, traditionally, to impose some sort of "order" on them. Stravinsky (1956) speaks of ". . . the need that we feel to bring order out of chaos, to extricate the straight line of our operation from the tangle of possibilities and the indecision of vague thoughts. . . . Now art is the contrary of chaos. . . . All art presupposes a work of selection. . . . To proceed by elimination—to know how to *discard,* as the gambler says, that is the great technique of selection."

The *degree* of selection differs among composers and among their compositions. One intuitively senses, for example, that Bach's music is more "ordered" than that of John Cage, one of the contemporary composers who intentionally introduce random elements and sequences of sound into the *cultural* sign system of their works. Any sequence is permissible according to Cage's structural principle of indeterminacy, while manifestly only patterns of a certain sort occur in Bach; Bach is more selective than Cage. The degree of selectivity of the works of composers is then a parameter of their style.

Styles as probability systems

If composers are Markov sources, that is, if their output may be considered to satisfy the assumptions of the theory, then the formulae for redundancy (Equation 3) measure the degree of order or structuration in their works. Hence, in order to justify the application of the redundancy measure, it must be shown that the sign system studied is actually a Markov chain. In support of this assumption is the description of musical styles as probability systems.

"Styles in music are basically complex systems of probability relationships in which the meaning of any term or series of terms depends upon its relationships with all other terms possible within the style system" (Meyer, 1956, p. 54).

For example, Frances Densmore (1918), in her many extensive studies of Indian music, described styles in statistical terms. She included tables listing the frequencies of tonalities, the relations of the first and last notes to the keynote, the number of ascending and descending intervals, the number of accidentals, the number of pieces recorded in various modes, the number of pieces beginning and ending with ascending or descending intervals, etc.; especially since many of the results were expressed in percentages is it apparent that the system is described as a probability system.

Not only may the musical score be characterized as a probability system, but the listener *perceives* the music, considered as a sequence of either cultural or psychological signs or both, as a probability system.

. . . the brain must, quietly, unobtrusively, incessantly, reckon the odds in favour of one event or one set of events implying another. In the simplest, isolated case, as in the laboratory, the odds are based on form—on the past history of that particular twin set of events being considered, on how they have come in before . . . (Grey Walter, 1953, p. 69).

This reckoning of probabilities applies immediately to music, where the odds *are* based on form; the brain automatically performs something similar to the syntactic studies to be surveyed.

In more intuitive terms, Hindemith (1952, p. 16) describes similarly the listener's act of perception.

While listening to the musical structure, as it unfolds before his ears, he is mentally constructing parallel to it and simultaneously with it a mirrored image. Registering the composition's components as they reach him he tries to match them with their corresponding parts of his mental construction. Or he merely surmises the composition's presumable course and compares it with the image of a musical structure which after a former experience he had stored away in his memory.

Relational perception in music

In recognizing the pattern represented by the score, the brain compares relations between successive events rather than absolute values. Instead of the absolute pitch of harmonies, the brain perceives relations between pitches, or intervals (McCulloch, 1949). Similarly, the brain perceives the pattern of relationships of a single melodic line.

. . . it is interesting to note that our recognition of the word or the tune depends mainly on the way the peaks of activity on the basilar membrane of the brain move to and fro, on their temporal pattern, in fact, rather than on their exact position in the line. When we hear the tune the absolute pitch of the sounds does not matter; it is not necessary that particular cells in the cortex should be excited, for what we recognize is a particular temporal sequence of activity in the auditory area (Adrian, 1947, p. 52).

Most applications of information theory to music are, therefore, in terms of pitch and time relations, rather than in terms of absolute values.

Methods of application

Three main approaches to the application of information theory to music have developed: analytic-synthetic, synthetic, and analytic.

Analytic-synthetic studies have used homogeneous bodies of existent music, such as nursery tunes, hymn tunes, and cowboy songs, to derive matrices of transition probabilities, then used the conditional probabilities to generate musical samples. These studies indicate the order of analysis, i.e., the size of the m-gram, necessary to produce musical samples like the original. The synthetic output indicates what characteristics of the original music the analysis took into account.

While there is no point in programming a computer to print out a Bach fugue when a Bach fugue is fed in, if the probabilities given a computer are considered analogous to the musical experience of a composer, and if the probabilities are based on a wide range of music, classical to modern, then the output might be compared with the current output of composers to see to what extent their works are probabilistic recombinations of their experience. (This has not yet been done.)

Synthetic works, studies in generating music, have in general followed Ashby's plan (1956) of using a random source to generate all possibilities and passing the random output through a selector. In a computer, formalized compositional "rules" may be translated into numerical terms and made to act selectively on a sequence of random numbers generated by the computer. It is then possible to see in musical terms exactly how much variety or constraint in the output the musical rules imply. Comparison of the machine's output with the output of a human given the same rules would show what elements of the human's output are uniquely human. The computer may be used to experiment with the effects on output of systematic variation of the restricting rules or of the introduction of entirely new selective rules. In fact, with arithmetic coding of rules, the computer may be used to mass-produce any kind of music.

Analytical works have primarily attempted to use redundancy as a parameter of style. The numerical value of redundancy has been proposed as a characteristic of a given body of music, e.g., the *Lieder* of the nineteenth century. In other cases, the analytical results were used to compare styles of different periods.

It has been suggested that analysis of a few "master-works" might yield optimal values for the information-content and redundancy which could be used as criteria to judge other compositions. There are two objections to this: first, syntactical analysis totally ignores external relations, such as audience sophistication and the point in musical history, which are important, if not dominant, in determining the "value" of a work; second, while a prejudice towards order or disorder may be helpful in practice, it is inappropriate in a theoretical analysis.

ANALYTIC-SYNTHETIC APPLICATIONS OF INFORMATION THEORY TO MUSIC

Cowboy songs

The first analytic-synthetic application of information theory to music was by Fred and Carolyn Attneave. Quastler (1955a) reported that the Attneaves analyzed Western cowboy songs to obtain transition probabilities for every note *preceding* a particular note. From a final note, C, they started a Markov chain with the proper probabilities going backwards, after selecting a standard form and rhythm. Two "perfectly convincing" cowboy songs resulted from a few dozen random walks of the chain, i.e., in a few dozen sequences generated according to the probabilities of the source. "It is legitimate to surmise that a slightly more elaborate statistical approach (possibly just extension to 'trigrams') would have resulted in a really good percentage of successful folk song compositions" (Quastler, 1955a, pp. 168-169).

Nursery tunes

The second application seems to have brought the possibility of using information theory in musical analysis to wide attention, because it is mentioned in most subsequent applications. Pinkerton (1956) transposed thirty-nine nursery tunes to the key of C. He calculated the probabilities of the seven tones of the diatonic scale plus the probability of a rest or hold.

On the basis of the simple probabilities, and considering the alphabet to have seven letters, Pinkerton computed a redundancy R_1 of 9 per cent. He tabulated the probabilities of pairs of notes and generated songs using a deck of twelve appropriately labeled cards as a random source.

Then he introduced rhythmic redundancy. A pause or hold was not permitted to begin a measure. Since the transitions from a given note varied with its position in the measure, Pinkerton constructed six transition matrices to account for each of the positions in a measure in 6/8 time. From each of the six, he selected either the most probable or the two most probable transitions, to construct a circular net of at most binary choices, which he called a "'banal tune-maker,' since it is hardly more inventive than a music box." (See Figure 1.)

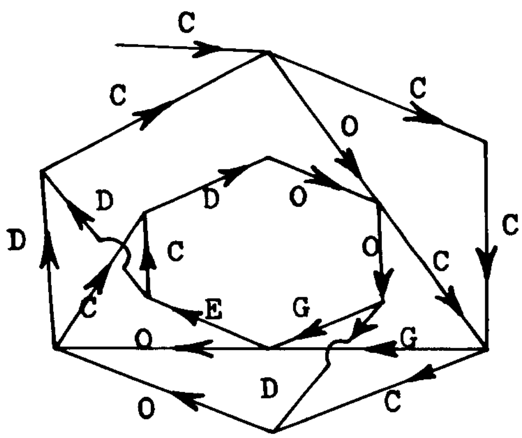

FIG. 1. Pinkerton's Banal Tune-Maker (Pinkerton, 1956, p. 78).

A sequence is generated on the banal tune-maker by following the line in the direction of the arrow and recording the note associated with each segment. Where there is a choice at a node, the toss of a coin (or any other method) may be used to decide which route to follow. Example 1 presents a sample generated by the tune-maker (p. 146).

On the basis of the thirty possible sequences of a measure's length and of the various probabilities, Pinkerton found the banal tune-maker to be more than 63 per cent redundant.

Teaching pieces

Sowa (1956) constructed "A Machine to Compose Music," using Pinkerton's techniques. The machine put in mechanical form a net similar to Pinkerton's but of greater complexity, which Sowa derived from his own analysis of piano-teaching pieces. Although the machine he used was actually slower than following the net on paper, the introduction of the computer opened the way for the time-consuming steps to be incorporated into the computer and the generating process speeded up. Example 2 is a sample of the output.

Hymn tunes

The most sophisticated work in analysis-synthesis is that of F. P. Brooks (Brooks *et al.*, 1957). The main points of his report are as follows:

Although present computers cannot induce generalized rules, but can only deduce, they are able to induce probabilities by counting relative frequencies and are able to use the results for probabilistic deduction.

Figure 2 presents the generalized schema for analytic-synthetic studies. The computer analyzes the given sample $S_1, S_2, \ldots S_n$ to obtain a set of probabilistic generalizations G. The output $S_1', S_2', \ldots$ is generated on the basis of the probabilistic rules in G.

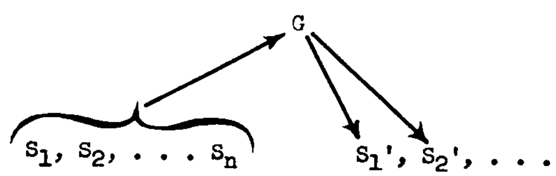

FIG. 2. General Model of the Analytic-Synthetic Process (Brooks *et al.*, 1957, p. 175).

Analysis of the sample consists of determining (1) the set of basic letters, the alphabet, and (2) the combinatorial relationships among them. The statement of combinatorial relationships may be (a) an explicit list of rules, (b) an exhaustive statement of all permissible combinations, or (c) a statement of the probabilities of the occurrence of letters and transitions between them. The last method is chosen: Markov analysis of order m determines the probability of each letter following each (m-1)-gram.

In Markov synthesis, an mth letter is chosen to follow a given (m-1)-gram according to the probabilities found in analysis. To choose according to probabilities, the Monte Carlo method (McCracken, 1955) is used (in this study and in all subsequent synthetic studies reported here). An example illustrates the method: to produce a sequence from the finite scheme

$$\begin{pmatrix} A & B & C \\ .22 & .75 & .03 \end{pmatrix}$$

we would generate or find a table of random numbers (a uniform distribution) running from 00 to 99. For all numbers from 00 to 21 we would write "A," from 22 to 96, "B," and from 97 to 99, "C" (a nonuniform distribution). A sufficiently long sequence reproduces the probabilities of the finite scheme.

Thirty-seven common-meter hymn tunes were chosen as the sample. Before analysis, their complexity was reduced by ignoring timbre and transposing them to the key of C. Silences in the single 4/4 lines were ignored and treated as holdovers from the previous notes, except for the beginning silence of each piece. No metric units shorter than an eighth-note were considered.

For treatment by the computer, the notes over a 4-octave range were represented by numbers. The different 8-grams were isolated and sorted in order within their 7-grams. The 7-grams were then sorted in order within their initial 6-grams, and so forth. For example, *AAAAAAAB* and *AAAAAAAC* are different 8-grams, but would be classed under the same 7-gram. Since each of the 8-measure hymns had sixty-four 8-grams, there were a potential $37 \times 64 = 2{,}368$ 8-grams. Only 1,701 distinct 8-grams were found; 1,531 distinct 7-grams were found.

In synthesis, 8-digit random numbers were generated. The first note was found from the probability tables of m-grams whose first m-1 notes were rests. The second note was found from the tables of m-grams whose first m-2 notes were rests and whose penultimate note had been chosen in the previous step; and so forth. The constraints imposed on the output were: (1) silence is ignored (as in analysis); (2) the first note of a measure must be struck; (3) each phrase must end with a dotted half-note; (4) each piece must end on C. If these constraints were not met by the first random number, up to fourteen others were tried; if no satisfactory one was found, the hymn was rejected.

Three types of metric constraints were used to give the hymns rhythmic redundancy.

Of some 6,000 hymns begun by the computer, 600 were completed. The table plotting per cent of completed hymns against order *m* of analysis-synthesis shows that the higher the level of restrictions the smaller the percentage of success.

Examples 3 through 7 illustrate synthetic results using Markov chains. When $m = 1$, i.e., when simple probabilities of occurrence are used (Example 3), the output is hard to sing and is definitely not hymn-like. The example contains the only accidental in the output; accidentals (sharps and flats) occurred less than one per cent of the time in the input sample. When $m = 2$ (Example 4), the output contains trigrams and progressions not in the input. The output for $m = 4$ (Example 5) is less rough; its extreme range is due to the original transpositions of the hymns. The synthesis reaches middle ground when $m = 6$, avoiding all three of the difficulties of analysis-synthesis (Example 6). When $m = 8$, some of the 8-grams are identical with those of the sample; the implicit structure of the 8-grams is strong enough to keep the output like the input even under the skeletal rhythmic constraints (Example 7). Hence eighth-order synthesis is too high.

Brooks concludes (1957, p. 182):

> Ideally, the present experiments would not have been needed. One would have described the information-content of the sample and the extent of the constraints of the musical structure, and from this one could have predicted that first and second-order Markoff analysis-synthesis would have yielded sequences unacceptable as hymns and that eighth order analysis-synthesis would have yielded sample members.
>
> The wide discrepancy between the ideal situation and that which currently prevails emphasizes the large amount of theoretical and experimental work which will have to be done before the inductive-deductive processes are well enough understood for general use in computing machine applications.

Composers' assistant

Olson and Belar (1961) have programmed a computer to "help a composer to create new music by suggesting variations and new tone combinations based on his original ideas."

The computer was given properly coded note sequences from Stephen Foster which reflect the composer's style, on the basis of which the computer selected new variations. This is simple analysis-synthesis.

The electronic coding of the newly generated music is converted to audible musical sounds in order to let the composer select what he likes. Hence the selectivity of the computer is supplemented by human selectivity to increase the redundancy of the output.

Generalizations

So far, analytic-synthetic studies have been exclusively concerned with melody, an abstraction from the cultural sign system of music, itself an abstraction. Because of the high level of abstraction, the techniques used have not been specifically musical in orientation: the notes which were coded by numbers could just as easily have been the colors of cars passing a toll station.

The melodic redundancy introduced by the note probabilities alone was insufficient to make the output comprehensible. While the output was properly an undifferentiated string of notes, since the input was no more than that, it was found necessary to introduce rhythmic redundancy through metric constraints.

Not all of the output was classed as "satisfactory." In effect, all the studies were more selective of output than the apparent synthetic mechanisms. Hence more constraint, i.e., higher redundancy, was required than the probability systems alone supplied.

Finally, it was not necessary to introduce the measures of information-content and redundancy, although Pinkerton (1956) and Olson and Belar (1961) did use them descriptively. It is possible to view a communicative process as a probability system, adopting some of the operational techniques of information theory, without applying the strictly limited measure of information-content.

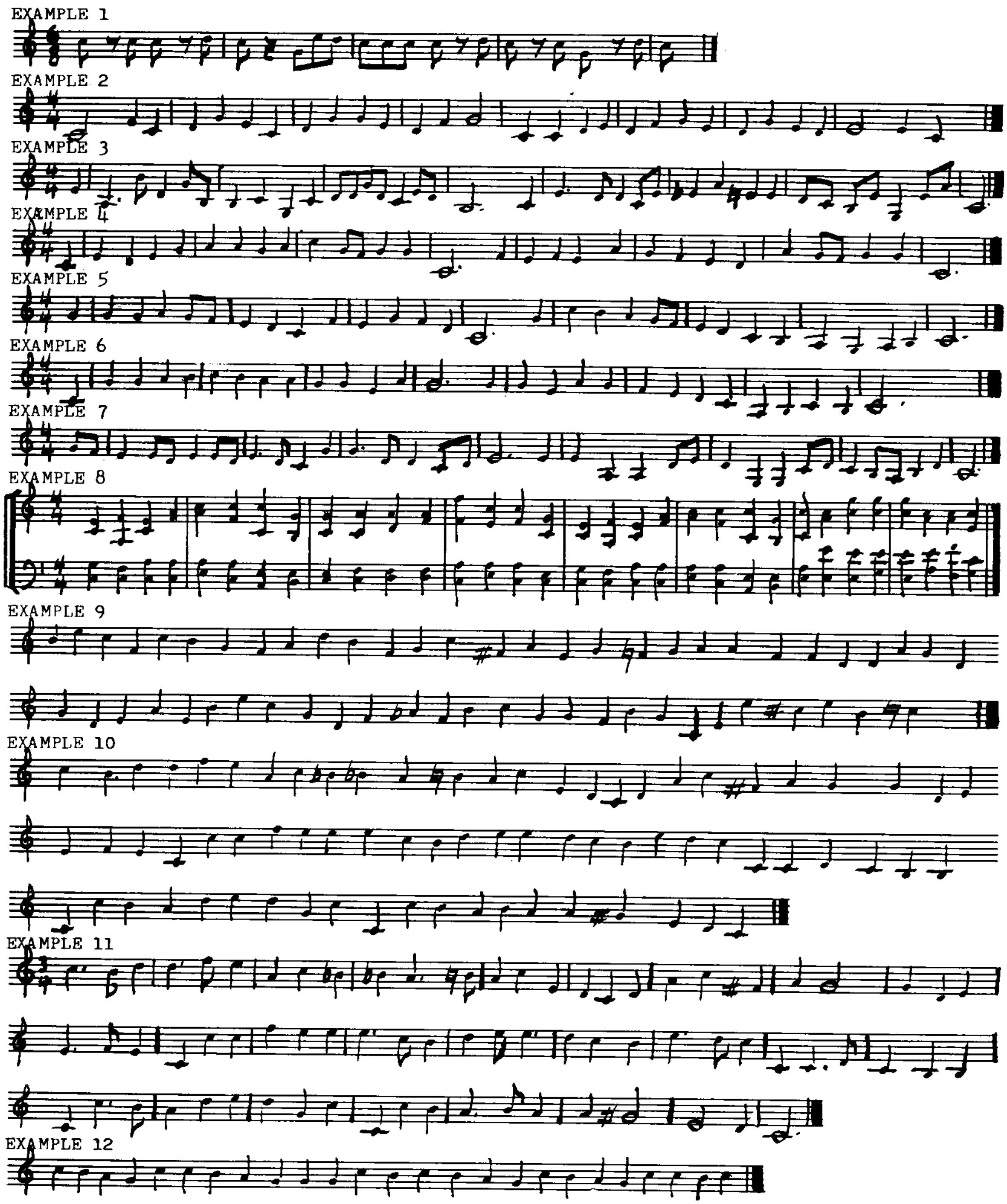

FIG. 3. Examples of Computer-Generated Music

SYNTHETIC APPLICATIONS OF INFORMATION THEORY TO MUSIC

Composition may be regarded as selecting acceptable sequences from a random source. Because such selection makes some sequences or classes of sequences more probable than others, composition implies, in the terms of information theory, imposing redundancy on a sequence.

Information vs. redundancy

A dipole is established. At one end are sequences with a relative information (relative uncertainty) of 1 and a redundancy of zero. At the other are sequences with zero relative information and 100 per cent redundancy. Sequences at the first pole are "noise": in the technical sense because any frequency of sound of any duration is as likely as any other, and in the musical sense because they have no "pattern," hence "make no sense." Sequences at the second pole are totally organized: the receiver has accounted for *m*-grams of a length such that no information is gained from additional output. Between the two extremes is a continuum of increasing redundancy and decreasing relative entropy.

While most musical literature falls between the poles, some approaches the extremes. Some examples show an appreciation of the variability of the order-disorder relationship.

Cage (1959) described the process of composition used in some recent piano music at the zero redundancy extreme. A master page is constructed on transparent plastic paper. Chance operations from the ancient Chinese *I-Ching* (Book of Changes) determine how many sounds are to occur per page. These are penciled on a transparent blank sheet. The location of the notes depends on the imperfections in the paper. The transparent sheet is put on the master sheet and note-heads are written where the penciled dots are. Whether the note-head is black or white depends on whether or not the dot falls on the staff. Eight tosses of a coin determine the clefs of the staves. Chance operations determine which notes are to be played normally, which played with the strings muted from inside the piano by the hand, and which plucked on the strings inside the piano. Similar chance operations determine whether a sharp, flat, or natural sign is prefixed to each note.

> These pieces constitute two groups of 16 pieces (21-36; 37-52) which may be played alone or together and with or without *Music for Piano 4-19*. Their length in time is free; there may or may not be silence between them; they may be overlapped. Given a programmed time-length, the pianists may make a calculation such that their concert will fill it. Duration of individual tones and dynamics are free (Cage, 1959, p. 43).

While the compositional methods of Cage are minimally selective, his output does not have zero redundancy, because it is restricted to the well-tempered scale available on staff paper and capable of performance on a piano.

At the other extreme, Cage has also composed a piece with a 100 per cent redundant internal structure, "4:44." It consists of four minutes and forty-four seconds of silence (Hiller, 1959, pp. 112-13). Since the set of possible events in "4:44" has only one letter (a "unit" of silence), the constant repetition of that letter conveys no additional information, and the redundancy is 100 per cent. (Considering the musical experience which most listeners bring to "4:44," however, such a long silence, as a sequence of musical events, is highly improbable; in a larger context, the piece is highly informative. See the critique below.)

The "totally organized" music produced by one contemporary school of composition should, theoretically, also be 100 per cent redundant (Rochberg, 1960). In theory, the whole of a composition derives from the constant repetition of a single permutation of the first twelve integers, applied to the twelve notes of the chromatic scale, the levels of dynamics, the note durations, etc. If the destination were able to code 12-grams for notes, dynamics, tempi, etc., totally organized compositions would be totally redundant, at least in the cultural sign system. However, the lack of redundancy on architectonic levels other than 12-grams usually makes them overly informative.

Synthetic experiments

Synthetic experiments make conscious use of information theory.

Hiller and Isaacson (1959, pp. 56-57) report that Caplin has coded the programming of Mozart's *Musikalisches Würfelspiel* for computer (von Köchel, 1947, p. 909).

Bolitho and Klein programmed a Datatron computer to write both the rhythm and the melody of popular song-tunes (Burroughs Corp., 1956). Although one tune generated by the Datatron, *Push-Button Bertha,* was given lyrics and broadcast, the project has been discontinued (Hiller & Isaacson, 1959, p. 56).

Minor projects in computer synthesis of music by Guttman and by Plyter (Hiller & Isaacson, 1959, p. 57; Anon., 1960) are notable in that the computers also performed the compositions.

Another experiment to compose music by creating a stochastic source was performed by Pierce (1956a) and Shannon. They made a catalogue of "permissible" chords on roots I through VI in the key of C and constructed three special dice. By throwing dice and using a table of random numbers, they chose sequences of chords. They imposed the restriction that two successive chords had to have a common tone in one voice.

They arbitrarily supplied a redundant rhythmic structure similar to that used in beginning harmony courses.

Three sequences of eight measures each were generated. Example 8 (Pierce, 1956a, p. 271) is the first of the three. Identifying characteristics of the sequences are the tendency of voices either to remain on one note or to jump wildly, and the consistent violation of the "rules" of musical theory. Evaluating, Pierce states (p. 273): "Acting in the capacity of a music critic, I find them pleasing rather than deep. They are less dull than poor hymns but are considerably inferior to Bach."

Pierce (1956b) also reported an experiment by Slepian. "Thus he had each of a group of men add to a 'composition' after examining only one or more preceding half measures."

Illiac Suite for String Quartet

The most sophisticated and extensive synthetic use of information theory to date is the series of experiments with the Illiac computer by Hiller and Isaacson (1959; Hiller, 1959; Anon., 1956; McKay, 1959; Livant, 1961). A sampling of the results of the experiments was published in four movements as the score to the *Illiac Suite for String Quartet.*

The first experiment showed that a computer can generate music in the complex but well-codified style of the counterpoint of Palestrina (1526-1594). In the experiment, the rules for strict first-species counterpoint were given to the computer. The Monte Carlo method was used to generate simple melodies, or *cantus firmi,* of from three to twelve notes in length.

Rules restricting the relations between notes sounding at the same time, that is, rules restricting "vertical" relations, were then introduced and two-part counterpoint was generated. A cadence subroutine (set of rules) was inserted. Samples of each length, from three to twelve notes, were generated.

The computer program was then extended to compose first species counterpoint for four voices. Although the $\begin{smallmatrix}6\\4\end{smallmatrix}$ chord and the VII_6 chord then became permissible, the increased interaction between the four voices led to a greater proportion of failures among sequences begun. A more complex cadence subroutine was introduced and samples of each length were generated.

For the second experiment, a new program was written embodying all fourteen rules of strict sixteenth-century counterpoint.

Numerous restrictions called for by the rules were imposed. "The total number of individual arithmetic instructions required by this program for writing strict counterpoint exceeded 1,900 individual operations" (Hiller & Isaacson, 1959, p. 110).

With this program, the computer generated "counterpoint of fair quality, strongly reminiscent, if one ignores a certain monotony in rhythm, of passages from Palestrina" (Hiller, 1959, p. 114). Rules were then successively withdrawn from the program until the computer was generating random "white note" music. The results of the experiment were assembled in reverse order. Thus the second movement of the *Illiac Suite* progresses from random diatonic music to music generated with the complete program.

In the third experiment, the computer's program was altered to introduce rhythmic, dynamic, and orchestral interest as well as a style of greater contemporary interest. Horizontal rhythmic redundancy and vertical redundancy were introduced by generating 4-digit binary numbers. "Here 0000 indicated that all four voices would be rhythmically independent, 1111 called for the same rhythm in all voices . . ." (Hiller, 1959, p. 117). Dynamics or loudness levels were coded. Random numbers were used to select *crescendo, diminuendo,* or no change in level. Vertical and horizontal redundancy was introduced for dynamics as for rhythm, but there was no correlation between the rhythmic and dynamic codes in the computer. The manner of performance was coded by numbering sixteen of the commonest playing techniques for stringed instruments and generating random numbers between zero and fifteen.

The notes in the third experiment were first selected simply at random. The chromatic, rather than the diatonic, scale was the alphabet. Complete lack of restrictions gave each note maximal information-content. Four restrictive rules were then coded.

Stylistically, "while the wholly random sections resembled the more extreme efforts of *avant garde* modern composers, the later, more redundant portions recalled passages from, say, a Bartok string quartet" (Hiller, 1959, p. 117).

Finally, tone rows according to Schönberg's formulations, interval rows, and a restricted form of tone row were generated and included as the coda to the third movement (Hiller & Isaacson, 1959, pp. 124-131).

In the fourth experiment a "new" formal principle was introduced to yield "Markov-chain music," i.e., a music in which the probabilities of successive notes depended on the preceding. In the first three experiments, the output was also Markov-chain music, but the probabilities were either zero (for a forbidden sequence) or equal for all permissible choices (random choice). In this experiment, two probability functions were assigned to the musical intervals. (1) The harmonic function increased the probabilities of those intervals which establish tonality. The unison was made most probable, the octave, second most probable, the interval of a fifth, third, and so on. (2) The proximity function made the unison the most probable, the minor second, second most probable, the major second, third, and so forth. A third function for interval probabilities was compiled by adding the harmonic and proximity functions for each interval. Zero- through third-order Markov-chain music was then generated. (Hiller and Isaacson were unfortunately mistaken as to the mathematical definition of the order of a Markov chain, so that what they describe as a second-order chain is actually a third-order chain.[2]) Finally, a simple closed structure was generated by making the Markov-chain probabilities effective on the strong beats of the measure. This increased the length of time over which the Markov-chain restrictions were effective. The sequences generated were for a single voice, but were transcribed into the four-voice score to make the score contain as much of the experimental results as possible.

Generalizations

The synthetic applications of information theory to music are really a special case of the general analytic-synthetic process (Fig. 2). In all synthetic experiments but the last two movements of the *Illiac Suite,* human analysis of existing musical samples yielded the body of rules and restrictions (G). In the last two movements of the *Illiac Suite,* the rules were arbitrarily assumed by the experimenters.

Corresponding to these two methods of arriving at the set of generalizations G are the two functions the synthetic output may serve. The generated sequence, whose source is a set of rules of known structure, may be compared with the samples of musical literature which were analyzed. Such a comparison would make it possible to determine, e.g., what part of a Palestrina motet may be said simply to follow formalized rules and what part may be said to result from the "human element."

[2] Compare Hiller & Isaacson (1959, pp. 142ff.) with Shannon (1957, p. 10).

When G is arbitrarily assumed, the output may be used to study what the rules of G mean in musical terms. Hiller and Isaacson (1959) have pointed out that one advantage of using a computer to generate output according to new rules is that the computer has no prejudice due to previous training. The computer is really no more selective than the rules given it.

The major result relating the degree of restriction of the rules (G) and the style of output is given by Hiller and Isaacson (1959, pp. 160-61):

> . . . simplicity of style and hence accessibility bears an *inverse* relationship to the freedom of choice. The simplest style requires the severest restrictions and has the highest degree of redundancy. On the other hand, simpler musical styles are by no means necessarily the easiest to write, since the difficulty of composition involved in making the best choices from among the many available in large structures in a less restricted style is offset by the fact that more of the available choices are permissible in terms of the desired effect.

The inverse relationship between simplicity of style and freedom of choice is verified by analytic studies.

ANALYTIC APPLICATIONS OF INFORMATION THEORY TO MUSIC

Mathematical analyses of music

Mathematical analyses of the cultural sign system of music date back at least to the French composer Rameau (1722). Setting forth his new harmonic theories, he said:

> Music is a science which ought to have certain rules; these rules ought to be derived from a self-evident principle; and this principle can scarcely be known to us without the help of mathematics. (Rameau, 1722, pp. 565-566).

In his mathematical discussion of symmetry, Weyl (1956) notes that Speiser tried to apply combinatorial principles of a mathematical nature to the formal problems of music. Speiser analyzed Beethoven's "Pastoral" Sonata for Piano, Op. 28; he also refers to the investigations by Lorenz on the formal structure of Richard Wagner's chief works. The primary mathematical tool used was apparently the theory of groups.

More recent analyses have been statistical Zipf (1949, p. 336) reported an analysis of the solo line of Mozart's Bassoon Concerto in B^{b} Major. He found a linear inverse relation between the frequency of an interval and its size, for ascending intervals, for descending intervals, and for both types combined. Smaller intervals were more probable than large intervals. The same distributions were in general found (p. 337) in the much shorter (797 notes) "Etude in F Minor," Op. 25, No. 2 by Chopin and in Irving Berlin's "Doing What Comes Naturally" and Jerome Kern's "Who." Such empirical results approximate the distribution of intervals given by the proximity function in the Markov-chain music of Hiller and Isaacson (Experiment 4). Quastler (1955a, p. 168) comments on Zipf's results:

> This is the distribution of free path lengths between random collisions; accordingly, one would produce the same distribution by alternating movements up and down and, on passing each key, interrogating a constant-probability random source whether or not to hit. This method will produce the right frequency distribution, but hardly the right music.

Fucks (1955) studied the formal structure (i.e., the cultural sign system) of the *Missa Papae Marcelli* by Palestrina and the *E-Minor Missa* by Bruckner. His first results include a function which gives different values for the works of Palestrina and Bruckner, and so may be useful as a stylistic criterion.

Albrecht (1956) states: "It is my hope ultimately to derive the now purely empirical laws of harmony and counterpoint as developed by Palestrina and Bach." Schillinger (1948) gives the results of statistical analyses in his *Mathematical Basis of the Arts.*

The analytic applications reviewed below gather data and apply the immediately given formulae of information theory.

Style as information

Youngblood[3] (1960) has attempted "to

[3] Youngblood's first report of his work (1958) is revised and extended in his thesis. Since the figures given in his thesis are the accurate ones (according to a letter from Dr. Youngblood, 15 February 1961), only they and the results based on them will be reviewed.

The unpublished theses of Youngblood and Brawexplore the usefulness of information theory as a method of identifying and defining musical styles." Musical style may be considered a probability system which, he argues, must be stationary: ". . . wholesale variations would be required before a musical style would become anything but stochastic [stationary], for it is this very homogeneity that makes it recognizable as a style" (p. 14). Information theory can be used to "measure the constraints under which various composers and groups of composers worked, and can furnish us with figures with which we can more accurately and more meaningfully describe these styles" (p. 18).

Youngblood selected a corpus of Romantic songs in major keys, consisting of eight songs from Schubert's *Die Schöne Müllerin,* six arias from Mendelssohn's *St. Paul,* and six songs from Schumann's *Frauen-Liebe und Leben.*

First, Youngblood found an approximation to the probability of each of the 12 tones of the scale for each composer. He calculated for each composer the first-order information-content using Equation 1 and the first-order redundancy ($\%R_1$), using Equation 3. From an approximation to the cumulative probabilities for the whole sample, he found the cumulative first-order redundancy.

Second, Youngblood found the frequency of each pair of tones for each composer and the redundancy of pairs of tones ($\%R(i, j)$). By Equation 2 he calculated the information gained from j, given i (H_2) and R_2 immediately followed by Equation 3. The redundancies found are as follows (p. 34):

TABLE 1
REDUNDANCY OF ROMANTIC MUSIC

Composer	$\% R_1$	$\% R(i,j)$	$\% R_2$
Schubert	12.5	20.4	35.6
Mendelssohn	14.9	24.0	43.5
Schumann	14.7	22.4	37.3
Cumulative	13.4	20.5	29.2

Mendelssohn's use of chromatic tones was much less frequent than Schumann's. Accordingly, Mendelssohn is more redundant. Schubert's comparatively small redundancy is due to modulations, rather than chromatic freedom.

> Even though they were virtual contemporaries, Schubert, Mendelssohn, and Schumann are different in style. That the figures in this chapter do not point up the differences clearly is less an impeachment of the technique than an indication that it is not in melody that they are so strikingly different (p. 42).

A harmonic analysis would probably bring out some of the differences (cf. the critique below).

To compare the data on Romantic music with those of another style, Youngblood analyzed the *Gloria, Sanctus,* and *Agnus Dei* from the First Mass for Solemn Feasts (*Liber Usualis,* pp. 19-22) and the *Kyrie* from the mass *Orbis Factor* (*Liber Usualis,* p. 46).

Considered as a 7-note system, that is, as a source whose alphabet has only seven letters, the Gregorian chant analyzed had a very low first-order redundancy; but the restrictions on relations between notes gave it a higher second-order redundancy, which, however, was still lower than that of any of the Romantic composers. Since for the contemporary listener there are theoretically twelve possible culturally significant tones at any given point, Youngblood calculated the redundancy of the chant as a 12-tone system. His results were (pp. 39, 42):

TABLE 2
REDUNDANCY OF GREGORIAN CHANT

Gregorian chant	$\% R_1$	$\% R(i,j)$	$\% R_2$
7-tone system	3.2	23.9	28.8
12-tone system	23.9	41.7	44.0

From the figures, it is apparent that the Gregorian chant as a 12-tone system is slightly more redundant than any one of the three Romantic composers, and considerably more than the three of them combined.

The value of this application is that it refines a sensibility to stylistic differences.

> Most musicians can at present either intuitively or on the basis of certain vague generalizations identify at least five or six historical styles. It seems, however, that it would be useful to find a means of identifying and quantifying the characteristic features of a style, as well as measuring the differences between styles, if for no other reason

ley are reviewed at length because of the difficulty that might be encountered in trying to obtain them.

than to provide a basis for understanding and evaluating contemporary music (Youngblood, 1958, p. 31).

The probabilities can be considered to determine a stochastic source approximating the structure of the source which generated the sample. To compare the structure of the source defined by the probabilities with the structure of the actual source, i.e., the composer, I applied the Monte Carlo method to the probabilities. I first constructed a "map" dividing the numbers from 000 to 999 according to the simple probabilities of tones found in Mendelssohn. Using a table of random numbers, I generated a sequence of 56 tones (Example 9) which is clearly too chaotic and too unpatterned to be Mendelssohn.

I then constructed 12 maps according to Mendelssohn's transition probabilities from each note and generated a sequence (Example 10). In Examples 9 and 10 I have supplied a few variations in octave. Example 10 bears no resemblance to Mendelssohn; but when rhythmic variety and metric pattern are supplied (Example 11), parts of it become quite tuneful.

To see the effect of imposing further redundancy on the probabilities of Mendelssohn, I arbitrarily selected the one or two most likely transitions from each note to construct a net similar to that of Pinkerton (1956). Example 12 is a sample of the net's output.

On the basis of analytic-synthetic results of Brooks (1957), it is plausible that analysis of at least twice the order performed by Youngblood (1960) is necessary to define a source which will generate acceptable Mendelssohnian tunes. On the other hand, Youngblood has pointed out that it may be the type, rather than the order, of analysis which needs to be changed. (He suggests as a possibility the technique used by Yngve [1956]. Cf. the comments below based on Chomsky.)

Information-content of rock and roll

Partly for comparison with Youngblood's figures and partly because of the intrinsic interest of the subject, I used Youngblood's techniques to analyze the melodic lines of "Hound Dog" and "Don't Be Cruel." The songs were transposed to the same key. Table 3 presents the frequencies of tones in each song and the combined frequencies.

TABLE 3
NOTE FREQUENCIES OF ROCK AND ROLL

Note	Hound Dog	Don't Be Cruel	Combined
I	40	39	79
II	0	0	0
III	40	18	58
IV	84	0	84
V	0	19	19
VI	0	0	0
VII	0	0	0
VIII	4	48	52
IX	0	0	0
X	4	83	87
XI	0	0	0
XII	0	17	17
Total	172	224	396

Table 4 presents the average information-content per note and the redundancy, as 7-tone and 12-tone systems, of each song and of the whole sample. Although the melodic line of "Hound Dog" uses a total of five different notes and that of "Don't Be Cruel" uses six, seven notes are used in the sample as a whole. Hence the redundancy of the sample is lower than the redundancy of either piece individually.

TABLE 4
INFORMATION-CONTENT AND REDUNDANCY OF ROCK AND ROLL

	Hound Dog	Don't Be Cruel	Combined
Information-content (bits/note)	1.73	2.32	2.61
Redundancy (% R_1)			
7-tone system	38.2	17.2	6.9
12-tone system	51.6	35.2	27.1

Incidentally, the amount of labor involved in this analysis and in the synthesis based on Youngblood's figures emphasizes the need for the assistance of a computer. The size of the sample analyzed is obviously too small to permit generalizations about the general body of rock and roll music, but the figures found are at least not at variance with what might be expected intuitively. Higher-order analysis would yield even greater redundancy, since the permissible sequences are very restricted.

Rhythmic style

Like the analysis of melody by Youngblood (1960), the analysis of rhythm by Brawley (1959, p. 3) is concerned "not so much with the intent of describing individual styles as with investigating the possibilities of a method for using information theory as a tool in describing styles." Two basic assumptions are made: musical rhythm is a discrete system of communication; musical style is an ergodic stochastic process which has the structure of a stationary Markov chain.

The main problem in rhythmic analysis is to decide what constitute the letters of this discrete system. ". . . in perception of occurrences in time, these occurrences are grouped into patterns" (Brawley, 1959, p. 19). Hence Brawley (1959, p. 20) concludes:

> Whether these groupings be regular or not, we shall adopt this grouping of any number of unaccented beats with a single accented beat as one of the basic principles of this study. More specifically, this will be the basis for determining the length of the primary, elementary symbols of our discrete system, i.e., the length of a single rhythmic unit, or rhythm.

Tempo plays an important role in determining rhythmic groupings. Rhythm is obviously produced by the muscles (rhythmic production is kinaesthetic, in other words), hence the tempo of rhythmic production is limited by the muscles; apparently kinaesthesis is also necessary to establish rhythmic perception. The "conscious present," on the order of one and a half seconds, is also a limit of the duration of rhythmic groups (p. 24).

As a solution to the problem of determining the limits of tempo within which one pulse is felt as one beat, Brawley adopts the range M. M. 60–120, i.e., a tempo with one or two beats per second. The limits of this range are flexible; e.g., M. M. 128 need not necessarily be changed to M. M. 64 (p. 26).

Brawley analyzed Bach's Two-Part Invention No. 14, commonly played at a tempo of *ca.* M. M. 68–72. He listed each rhythmic pattern in the invention, its frequency, and its relative frequency (probability). While both voices of the invention were taken into account, the upper part was generally given preference; the lower voice was analyzed when the upper part rested. The eight rhythmic patterns had an average information-content of 2.6 bits, and a redundancy $R_1 = 13.1\%$. The matrix of first-order transition probabilities was 15 per cent redundant.

Brawley observes, however, "it should be obvious that this analysis employing information theory is not very valuable" (p. 33). No generalizations are possible on the basis of this analysis because of the limited extent of the sample; hence a larger sample is needed.

A redundancy (presumably R_1) of 25.8 per cent was found for the three-part organum *Hec dies* in the style of Perotinus (early twelfth century). "One might, in fact, expect an even higher redundancy, but the use of the more prominate *[sic]* patterns in fairly equal numbers and the inclusion of a number of patterns only once or twice prevent higher redundancy" (p. 34). The development of rhythmic freedom by the time of Petrus de Cruce is reflected in a single motet of his, *Acun—Lonc tans—Annuntiantes,* which is only 15 per cent redundant.

Fourteenth-century polyphony had great rhythmic freedom. Analysis of two ballads of Matheus de Perusio gives a redundancy of 8.9 per cent.

Rhythmic analysis of the first movement of Schönberg's *Fourth String Quartet,* using performance markings as an aid to determining groupings, gives a redundancy of 6.4 per cent. But the analyzed sections of *Verklärte Nacht,* composed while Schönberg was still under Romantic influence, had a redundancy of 18.8 per cent.

For comparison with Youngblood's (1958) figures for melodic redundancy, Brawley (1959) analyzed three songs by Schubert and three by Mendelssohn. Schubert's songs were 5.7 per cent redundant and Mendelssohn's were 19.6 per cent redundant. "One of the principal reasons for the very low value found in the songs analyzed comes from textual demands *[sic]*" (p. 37). Analysis of three songs of Brahms gives a redundancy of 16.3 per cent; pooling the frequencies for all three composers gives a redundancy of 15.7 per cent, which possibly is a significant figure in the stylistic description of the nineteenth century *Lied*. "And the rather patent reason that the rhythm communicates in spite of the lack of rhythmic variety is that it is inevitably heard not simply in relation to itself, but in relation to the general style to which it belongs" (p. 39).

Finally, Brawley selected a large, closed body of literature for analysis, the minuets of Mozart's string quartets. The length of a measure (three beats) was taken as the length of the rhythmic group. Approximately one-third of the total number of measures was analyzed. The seventy-seven rhythms found (pp. 55-60) are 19.4 per cent redundant. The predominant use of a few rhythms negates the impression of variety given by this figure, however.

A matrix of transition probabilities for the twenty most frequent rhythmic patterns shows the tendency of several of these rhythms to repeat themselves immediately.

Latest studies

At the University of Illinois, there are several analytical applications of information theory to music in process.[4] Beau is completing an analysis of four sonata movements by Mozart, Beethoven, Hindemith, and Berg. By the fall of 1961, Fuller's analysis of Webern's Symphony, Op. 21, and Baker's studies of computer composition were to be finished. The last brings an informational theory of music full cycle back to analytical study of what is essentially analytical study in the first place.

Generalizations

Analytic studies, like synthetic studies, are a special case of the analytic-synthetic process. Their goal is to produce a set of generalizations characterizing the sample.

It was necessary first to establish an alphabet. The alphabets chosen so far have consisted of the minimal culturally-significant elements of music. Future studies will have to choose alphabets with letters of greater extent when the architectonic level analyzed is higher.

The analyses consisted in finding first- and second-order redundancies for various samples of stylistically homogeneous music. The aspects of the music considered were carefully abstracted from the whole piece; the score as a whole was never considered. The various figures for redundancy were compared and found to confirm pre-existing notions about the relative degrees of order of the samples analyzed. The finding of synthetic studies, that stylistic accessibility varies inversely as freedom of choice or information-content, was supported.

The authors of the analyses assumed that their figures for redundancy in some way described the listener's impression of the sample from which the data came. They assumed that the listener finds monotonous a piece with redundancy above a certain level, that he finds chaotic a piece with low redundancy.

CRITIQUE

For criticism, the information theory of music which has been reviewed may conveniently be divided into two parts (Figure 2). Analysis uses information theory to obtain a set of generalizations about a sample of music. Synthesis uses information theory to generate music on the basis of a set of generalizations or rules.

Criticism of the synthetic studies consists mainly of reviewing the output they have yielded. Criticism of the analytic studies consists mainly of reviewing the assumptions on which they are based.

Synthetic studies

Nothing of much musical value has resulted from synthetic experiments. This fact is less an indictment of the experiments than it is a recognition that their goal is to understand the mechanics of composition by creating models of the compositional process. The output should improve with improved models.

It may well be that information theory is essentially incapable of providing a model of the compositional process. Livant (1961) has pointed out that Chomsky's proof (1956) of the inadequacy of Markov sources to generate certain classes of sequences may apply directly to music. Chomsky shows that

[4] Letter from Dr. L. A. Hiller, 30 December 1960.

a Markov source cannot generate the class of all sequences in which the second half is the exact mirror image, or the exact repetition, of the first; in general, he shows that a Markov source is incapable of generating self-imbedded structures. Nonrhapsodic, or nonimprovisatory, music is rich in such self-imbedded structures; consider a fugue or a composition using serial technique.

The basic improvement required is greater interaction between the computer's output and the computer's program, i.e., a greater "feedback" from output to source. A composer does not apply the same sequences of tests or selective operations throughout a composition. His internal "program" is changed by his previous output. The computer is capable of adapting its program to its output by means of an order in the program to bypass sequences of orders under certain conditions. This bypass order or contingent programming has already been used in the programs of Hiller and Isaacson (1959). To generate a composition in a musical form with self-imbedded structures, such as a fugue or sonata, contingent programming will have to be much more frequently used.

Interaction between the parts of the program also needs to be increased. Rhythmic, performance, and melodic programs have so far been independent in computer programs, whereas in composing, rhythm, notes, and manner of performance affect each other. Different architectonic levels must be made to interact. Position in a phrase, and as Pinkerton (1956) observed, in a measure, affects individual note probabilities.

One valuable result of the synthetic studies is a very neat explanation of the listener's sensitivity to different styles. A note in the score was found to define a field of liberty. Analogously, the rules of a source define a field of liberty: a set of sequences of notes with a given redundancy-information relation. Just as the listener puts a set of psychological events into the same field of liberty, he puts a set of sequences defined by the same rules into a given field of liberty. The existence of such a way of listening is in fact apparent when we say that two pieces are the same in style, that is, are products of the same probability system.

The advantage of the synthetic studies over the analytic is that their working assumptions are necessarily explicit, because they must be incorporated into the computer program in order to affect the output.

Analytic applications

Many explicit and implicit assumptions lie behind the results of the analytic applications of information theory to music. The assumptions are of two general sorts. The mathematical assumptions assume that the musical sign system considered satisfies the requirements for the analytical application of the calculus of information theory. The aesthetic assumptions assume that the calculus of information theory is appropriate to use in saying something about music. The mathematical assumptions depend on the aesthetic, for if the calculus is inappropriate, there is no reason to make the mathematical assumptions in the first place. The aesthetic assumptions also depend on the mathematical, for if the mathematical criteria are not satisfied, the results of applying the calculus have no validity, no matter how "appropriate" information theory appears. Since the mathematical and aesthetic assumptions are thus interlocking, their division into two classes is more or less arbitrary and based on convenience.

Mathematical assumptions of analysis

Stochasticity. The first assumption for the application of information theory is that the source is stochastic, i.e., that no letter not already known to be in the alphabet can occur. For the cultural sign system, this assumption is easily fulfilled by letting every possible notation which has cultural significance be included in the alphabet. Then in any sample which does not use the whole alphabet, the unused letters have zero probability. The stochastic assumption is thus easily satisfied.

Ergodicity. The second assumption is that the source is ergodic, i.e., that a sufficiently large sample from an infinite sequence has the same statistical structure as the infinite sequence. This assumption implies that all sufficiently large samples from a given sequence have the same statistical structure.

To test this assumption, assume that there is an infinite sequence generated with the alphabet of the cultural sign system and that all music in existence, expressed in the cultural sign systems, constitutes the large sample. There is no contradiction to the assumption that the large sample has the same statistical structure as the infinite sequence, for the large sample is the only available evidence of the infinite sequence. This method of testing the ergodic assumption is, however, useless because it lumps all styles together into one probability system. One of the reasons for using information theory in the first place was to obtain a parameter of style.

Hence, to have a useful test of the ergodic assumption, the large sample must consist of a homogeneous corpus which is assumed to represent the style studied.

There is no evidence to indicate that the source of this corpus has a structure different from the corpus. The difficulty arises that any part of the corpus is not likely to have the same statistical structure as the whole. For instance, in Youngblood's analysis, each of the composers was a probability system different from the cumulated probability system. This difficulty is circumvented merely by saying that the works of any one of the composers do not constitute a "sufficiently large" sample.

But there is a circularity in this argument, for no means other than intuitive have been provided for determining whether or not a given corpus is stylistically homogeneous. There are at least two possible methods of determination. The first is arbitrarily to select a corpus of pieces or songs and define them to represent a large sample of a source. The obvious disadvantage of this method is that pieces whose statistical structures are widely disparate might become defined as output from the same source. The second method is to use information theory to achieve a parameter of style (say, per cent of redundancy) and let a given range of values of the parameter constitute homogeneity (e.g., variation of $\pm$ 5 per cent from a given value of redundancy). This second method would obviously be more successful in finding a parameter for what is recognized as a style, but it too is circular because it assumes that information theory may be applied to small samples, perhaps even individual pieces.

This last assumption demands that the small sample be stochastic, ergodic, and stationary (the last requirement has yet to be considered for the large sample). The small sample is obviously stochastic, by the same reasoning as for the large sample, but its ergodicity is doubtful. A small sample cannot contain two large samples, the statistical structures of which we can compare. Further, it is on arbitrarily large samples that probabilities depend. While it may be convenient to *assume* ergodicity in the small sample, the results based on the assumption then have no rigorous validity and must be considered grossly approximative.

The argument that the cultural sign system of music is ergodic may be summarized. While the entire body of existing music may be ergodic, the fact is useless because it yields no stylistic differentiation. Less inclusive but stylistically homogeneous corpora may be assumed to be ergodic; but an operational definition of their homogeneity would have the smaller samples selected either arbitrarily or on the basis of values of a stylistic parameter based on information theory. The latter operational definition of homogeneity assumes that small samples are ergodic, which cannot be established. The only other way operationally to define homogeneity in a sample is to find some nonintuitive stylistic parameter which is not based on information theory; but to do so defeats the goal of applying information theory in the first place.

The net result of this examination of the ergodic assumption is to expose the circularity of arguments, such as Youngblood's, that a musical style must be ergodic because its homogeneity makes it recognizable as a style.

We have barred intuitive criteria of homogeneity. If intuitive criteria are permitted, the ergodic assumption may be made; but arguments based on intuitive considerations do not lend much validity to the results of the theory.

Stationarity. A third assumption necessary to the application of information theory to a large musical sample is that the source is stationary, i.e., that the statistical structure of the sequence is independent of the time at which observation of the sequence begins. There is no evidence to refute this assumption, as before, if the whole body of music constitutes the large sample. If large corpora of stylistically homogeneous pieces of music are somehow determined, the stationarity assumption then assumes that within each corpus the structure of any large sequence is the same. This assumption obviously cannot be rigorously tested, because the corpus is too small to provide several "sufficiently large" sequences for comparison. For example, in a corpus such as the six partitas of Bach's *Clavier-Uebung*, the stylistic homogeneity of which is intuitively apparent, many of the sequences of sign-vehicles occur only once. No true estimate of relative frequency can be made on the basis of one occurrence. There simply is not enough music for the stationarity assumption to be tested.

Within a given piece, the stationarity assumption is not satisfied for two reasons. First, any single piece is much too short to give probabilities for a hypothetical source. Second, if one considers m-grams very much shorter than the whole piece, one can simply show that the relative frequencies change within the piece. For instance, in a movement in sonata-form, the first and last of the three sections (exposition and recapitulation) generally share the same statistical structure or are at least similar in structure, while the middle section, the development, is intuitively thought of as freer (less redundant) than the other two.

In general, because there is not enough homogeneous music to test it, the stationarity assumption rests insecure.

Markov consistency. A fourth assumption is that the source is consistent with regard to its Markov properties, that is, that there is the same order m of patterning throughout the sample. Chomsky (1956) has shown the difficulty of assuming a Markov source. Further, at the beginning of a piece, there can only be a small order of patterning in effect, while at the end, everything that has preceded affects, to a certain extent, the probabilities.

Infinite memory capacity. A fifth assumption is that the encoder of the sequence of musical events has an infinite memory capacity. The value of H, the number of binary digits required to encode a letter, is based on an optimal, most efficient code. Optimal coding requires the possibility of unlimited delay in encoding and decoding. Such unlimited delay is obviously not available.

Quastler (1955c, p. 27) examines the question: "What if a storage permitting infinite delay is not available?" Using the techniques of Fano (1949, 1950), Quastler finds that when the letters are equiprobable, "without any delay, one needs at worst 5.2% more than the minimum amount of code symbols per event. If one is allowed to code clusters of two events, one can always get within 3% of the optimum" (p. 28). In a binary choice with unequal probabilities, "a small delay allowing for coding 2–3 events at a time will suffice to yield nearly perfect efficiency . . ." (p. 29). In music, if the human brain is the encoder, the memory span prevents unlimited delay.

In summary, the mathematical assumptions of analysis are that the source is: stochastic, ergodic, stationary, and consistent with regard to its Markov properties; and that the encoder has infinite memory capacity. With the exception of the stochastic assumption, they are all difficult, if not impossible, to establish by analysis of a sample of music.

Aesthetic assumptions of analysis

Probability and expectations. Of the many aesthetic assumptions made in applying information theory to the analysis of musical style, the basic one is that statistical probability, or relative frequency, corresponds to the listener's expectations. Based on this assumption is the assumption that the *average* surprisal value H represents the listener's state of uncertainty while experiencing the sample analyzed. Youngblood (1960, p. 11), for example, assumes that information-content "refers to the degree to which the observer (hearer, reader) is in doubt as to what is going to happen next." Similarly, in rhythmic analysis, Brawley (1959, p. 17) assumes that he can use the calculus of information theory to "arrive at some more or less quantitative estimation of what the listener can expect, rhythmically, from a given style."

The probability system used in computing H for a given work or sample does not correspond to the listener's experience until he has stored the work or sample in his memory. The listener's experience is rather that of a constantly changing probability system. This probability system is based on his experience with the sample up to the present, rather than on over-all relative frequencies and average surprisals. The listener's feeling of uncertainty can only be measured by H if the listener has had a very long previous experience with musical sequences of exactly the same statistical structure as the sample analyzed. Unless the long previous sequence experienced by the listener has an average information-content of H, his feeling of uncertainty when experiencing the cultural sign system of a new sample will not be H, even though the sample when analyzed alone has an average information-content of H. Hence the assumption that the value of H of a corpus represents the listener's uncertainty while experiencing the corpus is, in general, not viable.

A calculus based on information theory, which quantifies the listener's experience of the changing probability system, has been developed by Kraehenbuehl and Coons (1959). In any sequence of differentiable events, the coming event in the sequence can either confirm predictions made on the basis of the past sequence, or fail to confirm ("nonconfirm") the predictions. "Information will be a measure of the degree to which a single prediction or an array of predictions is 'nonconfirmed' by the present event" (Coons & Kraehenbuehl, 1958, p. 129).

Predictions may be direct: e.g., if the past sequence is simply A, the only possible prediction for the next event on the basis of experience is A. Or predictions may be analogical: e.g., from the sequence of $ABCDEF$ the most plausible prediction is G.

On the basis of arrays of predictions, Kraehenbuehl and Coons (1959) develop an index of articulateness, representing average information flux so far in the sequence, and an index of hierarchy, representing average information reductions so far in the sequence. They assert that these two indices quantify and give operational meaning to the aesthetic terms "variety" and "unity." By finding the values of the indices for all possible 6-grams, they show that "a noticeable predominance of either results in a decrease in the effectiveness of both" (p. 129), i.e., that unity and variety are supplementary values rather than antagonistic.

The informational calculus of Kraehenbuehl and Coons (1959) rightly assumes that the listener cannot take account of what he has not yet experienced. It bases its quantifications of the listener's experience solely on the sequence immediately in process.

Yet the listener usually brings some previous experience, an averaged probability system, to the present sequence. This previous experience alters the probabilities and the associated informational values of the events in the sequence in process. Because it takes no account of previous experience, the calculus of Kraehenbuehl and Coons (1959) is inadequate.

What is required is a theory between the "traditional" information theory and the modified form created by Kraehenbuehl and Coons: one that accounts for the past experience of the listener in generalized, averaged terms, while measuring the current information flux of the present musical experience.

Interaction among aspects of the sample. Another aesthetic assumption is that one portion of the cultural sign system can be legitimately abstracted from the whole, and that values based on this abstraction will have the same worth as when the portion is a part of the whole.

Attacking this assumption, Albrecht (1956) writes:

> While I started along lines similar to Dr. Pinkerton's, I soon found that the harmonic context of a

> note was of primary importance in assessing its statistical value. . . . Thus mere concern over the frequency of occurrence of certain melody notes without reference to the harmonic context can produce meaningful results only in the trivial case where the entire melody is accompanied by the same tonic triad.

Similarly, the low value for rhythmic redundancy in Schubert's melodies (Brawley, 1959) had to be explained by the demands of the text. In general:

> An estimate of a probability which is made simply on the basis of unanalyzed samples or trials is not likely to be a safe basis for prediction. If nothing is known concerning the mechanisms of the situation under investigation, the relative frequencies obtained from samples may be poor guides to the character of the indefinitely large population from which they are drawn (Nagel, 1955, p. 401).

An understanding of the mechanism of interaction among the aspects of the musical sample is necessary to the proper analysis of each aspect.

Interaction among levels. A further assumption of the theory, at least as it has been applied, is that the sequence of musical events is experienced on only one architectonic level: in melodic analyses, on the level of notes or intervals; in rhythmic analyses, on the level of the pulse pattern. Obviously this assumption is inadequate for a picture of musical experience, because music is perceived simultaneously at many levels. In the cultural sign system of melody, for example, the notes, intervals, motifs, phrases, periods, sections, and movements all constitute letters of different alphabets. The theory will have to take account of the interaction among the levels.

Generalizations

Examination of some of the many working assumptions of informational analysis shows that many refinements will be required before the results can be taken seriously. Many of the numerical results given to three significant figures (Brawley [1959] even gives one result to seven figures) are no more than very rough approximations, if they have any meaning at all.

Future analytical applications will have to be much more carefully established. They will have to improve and adapt their mathematical tools. They will have to achieve a greater sensitivity to, and understanding of, the uniquely musical nature of musical experience. The observation (Tischler, 1956) that the internal relations of a work of art are unique to the medium demands an analytical tool sensitive to uniquely musical problems. The analytical beginnings made with information theory are a great advance over some of the operationally meaningless aesthetics, but a refined analytical technique is still distant.

PROSPECTS

Several questions are now being asked of the information theory of music. What can an information theory of music say about the emotional or affective response to music? What can it say about the psychological sign system of music? How is it related to the general problem-solving of human activity? What, in particular, are its effects on the hoary domain of aesthetics?

A review of partial answers to these questions may indicate some of the prospects for an information theory of music.

Emotion and music

There are two major theories to explain emotion in music in terms of information theory.

Meyer (1956) starts from the basic psychological theory that frustration or inhibition in the fulfillment of expectations usually lead to affective response. A musical style is a system of expectations. When these expectations are not fulfilled, e.g., when an improbable event occurs or expectation is inhibited by the ambiguity of the situation (Dibner, 1958), the listener may respond affectively with "musical emotion" or rationalize his frustration into "musical meaning." The response depends on his orientation. Hence "it would seem that the psycho-stylistic conditions which give rise to musical meaning, whether affective or intellectual, are the same as those which communicate information" (Meyer, 1957, p. 412).

Meyer (1957, p. 422) reviews some of the requirements for the statistical analysis of style. Note that they have all been ignored in practice.

1. The samples collected must take account of the tendency of systemic uncertainty to diminish and of designed uncertainty to be introduced as the music unfolds. . . .

2. Tonal probabilities exist not only within phrases and smaller parts of a musical structure but also between them. These probabilities are not necessarily the same. . . . Different sets of probabilities must be discovered for different architectonic levels.

3. It is a mistake to suppose that probability remains relatively constant throughout musical works. . . . Subsystems must be analyzed within the larger probability system. . . .

4. In defining the limits of a sample and discussing the probabilities involved, it is important to be cognizant of the historical development of musical styles. . . .

5. Not all the probabilities embodied in a musical composition are determined by frequency. . . . One of the preliminaries to a statistical analysis of musical styles must be a description and analysis of the constants involved in the psychology of thought.

He concludes by suggesting a general relation between information-content, musical meaning or emotion, and musical value.

The second informational theory of musical emotion has been developed by Moles (1956, 1958). He establishes first that the listener's psychological structure determines the logical structure of the message (cf. p. 138). He shows that the musical message may be interpreted simultaneously as various letters from various alphabets, depending on the scale of time perception. Moles distinguishes the cultural sign system (the "semantic" message) from the psychological sign system (the "aesthetic" message). The "aesthetic" message is unique to a given performance of a "semantic" message; now that recordings exist the "aesthetic" message may be studied. After a detailed examination of the psychological sign system, Moles (1958, Ch. 5) shows that while the "semantic" message is largely redundant, the information-content of the "aesthetic" message overwhelms the listener and creates the emotional response.

The theories of Meyer (1956, 1957) and of Moles (1956, 1958) are in no way antithetical. Meyer's applies to the cultural sign system to which the musician is usually trained to respond, while Moles' applies to the psychological sign system to which primarily the nonmusician responds. Both are essential, because there is a very serious weakness in studying only the cultural sign system:

> Though a specific statistical structure may be ascertained in a musical composition from an examination of the score, it is by no means certain that this structure, conceived at an intellectual level, is automatically transferred into sound (Meyer-Eppler, 1958, p. 58).

Hence studies of the psychological sign system are essential.

The psychological sign system

Much of the research applying information theory to psychology pertains directly to an informational theory of the psychological sign system of music.

Quastler (1955b) has shown that a human being's maximal rate of information transmission is around twenty-five bits per second. The limiting rate of transmission of professional pianists was around twenty-two bits per second.

Jacobson (1951) found that while the ear can transmit a maximum of ten thousand bits per second, in musical listening there are a maximum of ten independently perceptible notes per second. Each note represents a choice from the approximately one hundred semitones in the instrumental range, "for the musically trained ear."

> This gives about 70 bits/sec in the note perception, which accounts for the major portion of the informational content of the music. Thus it is evident that the brain can digest generally less than 1 per cent of the information our ears will pass (p. 471).

Hence the psychological sign system conveys at least one hundred times more information than the cultural, and Moles' theory that the listener is simply overwhelmed is supported.

Miller (1960) has studied the performance curves of individuals with an overloaded input of visual information and the relative importance of the various adjustment mechanisms they used. Similar experiments with aural information overload might be of assistance in understanding, for example, hostile reactions to overly informative music.

Attneave (1959) reviews other informational research on the psychology of hearing. Pollack (1952) found that subjects' capacity for information transmission through the absolute identification of tones is $\log_2 5 = 2.3$ bits, i.e., that in a scale of tones, subjects could *identify* only five correctly. Attneave suggests that a diatonic scale with a subjective standard for a musically trained subject to remember might transmit more information. In support of this, Rogers found that a concertmaster of a symphony orchestra could transmit 5.5 bits, i.e., could identify correctly on an absolute scale about forty-six tones. The concertmaster had absolute pitch (Attneave, 1959, p. 71). Psychological studies will ultimately have to explain the effects on musical perception of background and training.

Heuristic programming and music composition

Music composition is selection. Problem-solving is also selection of one out of a set of possibilities. Reitman (1960) has related musical composition to general problem-solving activities through heuristic programming of computers.

Heuristic programs are "attempts to incorporate in computer programs processes analogous to those used by humans in dealing intelligently with ill-structured problems" (Reitman, 1960, p. 410). Ill-structured problems include the discovery of theorems and proofs in mathematics and logic, the distribution of labor over jobs in manufacturing, the formation of scientific hypotheses in laboratory experiments, the selection of moves in chess, and the composition of music.

> One can invent a hierarchy of complexity which runs from elementary symbolic logic, to chess, and then to music. The solution of a logic problem is defined by the theorem to be proved. . . . Artistic works, however, may well have to satisfy networks of tests which themselves change as work progresses. (Reitman, 1960, p. 413).

To see how far heuristic programming can be extended in solving such problems, Reitman and Sanchez (Reitman, 1960) have tried to create a program simulating the behavior of a composer writing a fugue. They found a composer willing to try to describe his working methods and recorded everything he said or played over a period of several months. Their analysis of the data made it apparent that the computer would first have to be taught the elements of music. They programmed the computer with an information-processing language which enabled the computer to compose simple melody, harmony, and counterpoint. To solve the original problem of composition, they then incorporated an already existing General Problem Solver program (Newell, Shaw, & Simon, 1958). In so doing they are finding out just how general the General Problem Solver is.

Information theory and aesthetics

Perhaps the main value in the application of information theory to music lies not in the collection of specific results obtained, but in the mental habits developed when dealing with an aesthetic object or process such as music.

One trouble with many aesthetic discussions and treatises is that unless the reader already knows what the author is trying to say, the author may completely fail to communicate any message. Discussions of music as a "significant form" (Langer, 1952) and such statements as, "A melody should make sense and express a complete self-sufficient meaning" (Edwards, 1956, p. xxiii) are not particularly enlightening.

In the information theory of music, the concepts and statements are in strictly operational form; they designate actual concrete operations. Instead of defining "order" in terms of some ephemeral sense of "aesthetic pleasure" which not everyone may share, a new aesthetic based on information theory can speak of the redundancy of a given source in terms of universally reproducible operations. Instead of verbalizing about "unity" and "variety" in a series of fine distinctions, one can simply calculate the value of an "index of hierarchy" and of an "index of articulateness" or of other more adequate measures. Even specifically musical emotion becomes operationally measurable. Meyer (1956, p. 32) notes that musical affect may be objectively studied through the sequences of stimuli which give rise to it: ". . . granted listeners who have developed reaction patterns appropriate to the work in question, the structure of the affective response to a piece of music can be studied by examining the music itself." The information theory of music provides one means of examining the music itself. Certainly it is not the only means, nor even, at present, a particularly sensitive means, however.

Information theory alone cannot say what the nature of musical experience is. It can only be applied to a conception of musical experience which is arrived at by other means. To refine the conception of musical experience is a goal of aesthetics.

REFERENCES

Adrian, E. D. *Physical background of perception.* London: Oxford Univ. Press, 1947.

Anonymous. By the numbers. *Musical Amer.,* 1956, 76, 11, 13.

Anonymous. Why computers take up games. *Business Week,* Nov. 26, 1960, 138.

Albrecht, G. Letter, *Sci. Amer.,* 1956, 194, 4, 18-19.

Ashby, W. R. Design for an intelligence-amplifier. In C. E. Shannon & J. McCarthy (Eds.), *Automata studies.* Princeton: Princeton Univ. Press, 1956. Pp. 215-234.

Attneave, F. *Applications of information theory to psychology.* New York: Holt, 1959.

Brawley, J. G., Jr. Application of information theory to musical rhythm. Unpublished master's dissertation, Indiana Univ., 1959.

Brooks, F. P., Jr., *et al.* An experiment in musical composition. *IRE Trans. Electronic Computers,* 1957, EC-6, 175; correction, EC-7, 60.

Burroughs Corporation. Syncopation by automation. *Data from ElectroData,* August, 1956.

Cage, J. To describe the process of composition used in 'Music for Piano 21-52.' *Die Reihe* No. 3: *Musical Craftsmanship,* 1959, 41-43.

Chomsky, N. Three models for the description of language. *IRE Trans. Information Theory,* 1956, IT-2, 3, 113-24.

Coons, E., & Kraehenbuehl, D. Information as a measure of structure in music. *J. Music Theory,* 1958, 2, 127-151.

Densmore, F. *Teton Sioux music.* Bureau Amer. Ethnol. Bull. 61. Washington: Smithsonian Institution, 1918.

Dibner, A. S. Ambiguity and anxiety. *J. abn. soc. Psychol.,* 1958, 56, 165.

Edwards, A. C. *The art of melody.* New York: Philosophical Library, 1956.

Fano, R. M. The transmission of information. M.I.T. Res. Lab. Elect. Tech. Rept. 65 (1949) and 149 (1950).

Fucks, W. Reply, following Mathematical theory of word formation. In E. C. Cherry (Ed.), *Information Theory—Third London Symposium.* New York: Academic Press, 1955, p. 169.

Grey Walter, W. *The living brain.* New York: Norton, 1953.

Helmholtz, H. *On the sensations of tone.* Trans. A. J. Ellis. New York: Dover, 1954 (reprint).

Hiller, L. A., Jr. Computer music. *Sci. Amer.,* 1959, 201, 112.

Hiller, L. A., Jr., & Isaacson, L. M. *Experimental music.* New York: McGraw-Hill, 1959.

Hindemith, P. *A composer's world.* Cambridge, Mass.: Harvard Univ. Press, 1952.

Jacobson, H. Information and the human ear. *J. acoust. Soc. Amer.,* 1951, 23, 463-471.

Khinchin, A. I. *Mathematical foundations of information theory.* Trans. R. A. Silverman & M. D. Friedman. New York: Dover, 1957.

Kraehenbuehl, D. & Coons, E. Information as a measure of the experience of music. *J. Aesthetics & Art Criticism,* 1959, 17, 510.

Langer, S. K. *Philosophy in a new key.* (2nd ed.) Cambridge, Mass.: Harvard Univ. Press, 1952.

Livant, W. P. Review of *Experimental music,* by Hiller & Isaacson. *Behav. Sci.,* 1961, 6, 159-60.

McCracken, D. Monte Carlo method. *Sci. Amer.,* 1955, 192, 5, 90.

McCulloch, W. S. Brain as a computing machine. *Electrical Engineering,* 1949, 68, 492-497.

MacKay, D. M. Place of 'meaning' in the theory of information. In E. C. Cherry (Ed.), *Information Theory—Third London Symposium.* New York: Academic Press, 1955. Pp. 215-225.

McKay, G. Review of *Experimental music,* by Hiller & Isaacson. *J. Research in Music Ed.,* 1959, 7, 232.

Meyer, L. B. *Emotion and meaning in music.* Chicago: Univ. of Chicago Press, 1956.

Meyer, L. B. Meaning in music and information theory. *J. Aesthetics & Art Criticism,* 1957, 15, 412.

Meyer-Eppler, W. Statistic and psychologic problems of sound. *Die Reihe* No. 1: *Electronic Music,* 1958, 55-61.

Miller, J. G. Information input overload and psychopathology. *Am. J. Psychiat.,* 1960, 116, 8, 695-703.

Moles, A. Informationstheorie der Musik. *Nachrichtentechnische fachberichte,* Band 3: *Informationstheorie,* 1956, 47-55.

Moles, A. *Théorie de l'information et perception esthétique.* Paris: Flammarion, 1958.

Morris, C. W. Foundations of the theory of signs. *International Encyclopedia of Unified Science,* Vol. 1, No. 2. Chicago: Univ. of Chicago Press, 1955.

Nagel, E. Principles of the theory of probability. *International Encyclopedia of Unified Science,* Vol. 1, No. 6. Chicago: Univ. of Chicago Press, 1955.

Newell, A., Shaw, J. C., & Simon, H. A. Report on a general problem-solving program. The RAND Corp., P-1584, Dec. 1958.

Olson, H. F., & Belar, H. Aid to music composition with a random-probability system. *Science,* 1961, 133, 3461, 1368. (Abstract)

Pierce, J. R. *Electrons, waves, and messages.* Garden City, N.Y.: Hanover House, 1956. (a)

Pierce, J. R. Letter. *Sci. Amer.,* 1956, 194, 4, 18. (b).

Pinkerton, R. C. Information theory and melody. *Sci. Amer.,* 1956, 194, 2, 77.

Pollack, I. Information of elementary auditory displays. *J. acoust. Soc. Amer.,* 1952, 24, 745.

Quastler, H. Discussion, following Mathematical theory of word formation, by W. Fucks. In E. C. Cherry (Ed.), *Information Theory—Third London Symposium.* New York: Academic Press, 1955. P. 168. (a).

Quastler, H. Studies of human channel capacity. In E. C. Cherry (Ed.), *Information Theory—Third London Symposium.* New York: Academic Press, 1955. Pp. 361-71. (b).

Quastler, H. Editorial. In H. Quastler (Ed.), *Information theory in psychology.* Glencoe, Ill.: Free Press, 1955. (c)

Rameau, *Traité de l'harmonie* (1722). Extracts in O. Strunk (Ed.), *Source readings in music history.* New York: Norton, 1950.

Reitman, W. R. Information processing languages and heuristic programming. *Bionics symposium* (WADD Tech. Rep. 60-600). Wright-Patterson Air Force Base, Ohio: Directorate of Advanced Systems Technology, 1960, p. 410.

Rochberg, G. Indeterminacy in the new music. *Score,* January, 1960.

Schillinger, J. *Mathematical basis of the arts.* New York: Philosophical Library, 1948.

Shannon, C. E. A mathematical theory of communication. Bell Tel. Labs. Monograph B-1598, 1957. (Reprint of *Bell Syst. Tech. J.,* 1948, 27, 379-423, 623-656.)

Sowa, J. A machine to compose music. New York: Oliver Garfield Co., Inc., 1956.

Springer, G. P. Language and music: some parallels and divergencies. In *For Roman Jakobson.* The Hague: Mouton, 1956, p. 504.

Stravinsky, I. *Poetics of music in the form of six lessons.* New York: Vintage Books, 1956.

Tischler, H. The aesthetic experience. *Music Review,* 1956, 17, 189.

von Köchel, L. R. *Chronologisch-thematisches Verzeichnis sämtlicher Tonwerke Wolfgang Amade Mozarts.* Ann Arbor: J. W. Edwards, 1947.

Weaver, W. The mathematics of information. In *Automatic control.* ("A *Scientific American* Book") New York: Simon & Schuster, 1955.

Weyl, H. Symmetry. In J. R. Newman (Ed.), *The world of mathematics.* New York: Simon & Schuster, 1956. Pp. 1, 703.

Yngve, V. H. Gap analysis and syntax. *IRE Trans. Information Theory,* 1956 IT-2, 3, 106-112.

Youngblood, J. E. Style as information. *J. Music Theory,* 1958, 2, 24.

Youngblood, J. E. Music and language: some related analytical techniques. Unpublished doctoral dissertation, Indiana University, 1960.

Zipf, G. K. *Human behavior and the principle of least effort.* Cambridge, Mass.: Addison-Wesley, 1949.

(Manuscript received June 21, 1961)

An error appeared in Figures 4 and 6 of the article "Three-Person Non-Zero-Sum Nonnegotiable Games" by Anatol Rapoport, Albert Chammah, John Dwyer, and John Gyr, which was published in the January, 1962 issue of *Behavioral Science*. In these figures the point labeled 1 should have co-ordinates —18, 20; the point labeled 5 should have co-ordinates —14, 27; the point labeled 9, co-ordinates —16, 25.

The correct positions for points 1 and 5 are shown in Figure 2.