

---

## The perception of musical rhythm and metre

---

Mark J Steedman

Department of Psychology, University of Warwick, Coventry CV4 7AL, England

Received 2 February 1977, in revised form 13 March 1977

---

**Abstract.** The occurrence of relatively long notes, and the repetition of melodic phrases are important cues to the metre, or regular beat, of a piece of music. A model of how people use this information to infer the metre of unaccompanied melodies is described here. The model is in the form of a computer program, and involves a definition of melodic repetition which encompasses repetitions that include certain kinds of variation. The program has been applied to the task of analysing the metric structure of the forty-eight fugue subjects of the *Well-Tempered Clavier* by J S Bach. The program is discussed in relation to other models both of musical understanding and of sequential concept learning.

### Introduction

When people beat time to a melody, they perform an act of musical understanding analogous to the understanding of an utterance of their native language. In the sound they hear, the notes that fall on the beat are not explicitly marked, any more than there is explicit indication of the structure of a sentence in a speech wave. Moreover, neither the melody nor the sentence need have been encountered previously. Totally novel examples of either can readily be understood by almost anyone.

The beat is only one level of the metre, or regular temporal structure of a piece. A room full of people dancing to a waltz will not only show their agreement as to which notes fall on the beat. They will also show in their dancing that they understand the beats to be grouped in threes, and be similarly unanimous as to where the first, accented, beat of every group occurs.

Traditional musical notation expresses this simple aspect of the meaning of a piece in its written form, the score. For example, in figure 1b the notes in each main group of beats, or bar, are separated from the next one by a solid vertical bar line. Similarly, the short notes within each of the main beats of the bar (here indicated with dotted bar lines) are beamed together, making it clear that there are two beats to the bar and that each of these is further subdivided in a certain way. The position of the bar lines also identifies the notes on which the first, accented beat of the metre falls. This is exactly the kind of description of the piece that is necessary for someone to dance or beat time to it. Thus a theory of how one could get from the notes of a performance to the score that reflects these simple aspects of musical understanding would be of quite general interest to psychologists of music, and of cognition in general. Although the inference of metre is a very lowly aspect of musical understanding, there has been no complete account of how even this much understanding is achieved.

Recently, Longuet-Higgins (1976) has described a program which models several aspects of this kind of musical understanding. His program transcribes a live performance of a tonal melody into the equivalent of standard musical notation. One part of the program's task is to indicate the metric structure of the melody—the position of the bar lines with respect to the notes, and the subdivisions of the bar into successively smaller units. The problem for the program is that in real performance a musician may diverge considerably from a strict adherence to the

metre, in the interest of an expressive performance. The performer may gradually speed up or slow down, changing the absolute duration of the beat, and may make considerable variations in the durations of its subdivisions. In order to cope with this, Longuet-Higgins' program must at present be supplied with the absolute position of at least two principal beats at the very start of the piece—from then on it can happily deal with a wide range of expressive deviations from the initial metre, adjusting its representation of the metric structure as it goes along.

The initial metre is set up by requiring the performer to preface the melody by a bar's worth of beats on some low note. However, human beings are able to achieve this understanding without such help—they seem to be able to get their metric bearings as they go along, even when the melody begins on an offbeat, and not at the beginning of a metric unit. This paper is addressed to the question of how they do this, a problem that may be viewed as complementary to Longuet-Higgins' question of how, having done this much, they *keep* their bearings in the face of the vagaries of expressive performance.

For a single unaccompanied line of melody, it may be trivially easy to identify the position of the bar lines if the performance is so expressive as to make the first beat of every bar louder than the others. However, people can usually manage without such dynamic cues, and the inference of metre from 'deadpan' performances of unaccompanied melodies constitutes a well-defined and nontrivial task of musical understanding.

Longuet-Higgins and Steedman (1971) reported an early model of metrical inference and of the related problem of inferring key. In the next sections this and other earlier work are reviewed, and an extension to the preliminary model is described.

## 2 The problem

The problem for a model of people's performance on these tasks is that deviations from the key and metre are commonplace, even in the simplest melodies. A pitch which is not among the set associated with the key, called an 'accidental', may nevertheless occur in a piece. Similarly, although the metre defines the places where accented notes tend to occur, nevertheless under certain circumstances a note which is accented, say by being relatively long or at the beginning of a repeated melodic phrase, may occur other than on the first beat of a metric group. Such an occurrence is called a 'syncopation', and there is an example in fugue 2 of the first book of the *Well-Tempered Clavier* (see figure 1a). In this melody the longest note of all (which is indeed perceived as accented) falls on a weak, unaccented beat. (The note is marked with a star in the figure.) Nevertheless, this melody is metrically one of the most unambiguous of all the forty-eight. Even in the most unexpressive performance, it is hard to believe that the note in question could be perceived as anything but a syncopation, inconsistent with the metre. (The metre is so clear that it emerges even if the sequence of durations in figure 1a is merely tapped out with a pencil.) The syncopation does not mislead the listener into thinking that the metre of this piece is a quarter-note later in its phase than the one that Bach indicates with his bar lines. The accented note is comparatively late in the piece, and it seems that the metre has already been established in the listener's mind well before its arrival.

This suggests that a procedure for deciding the metric structure of melodies must place great reliance upon very early accents, and refuse to let this evidence be overthrown by later accents, such as the one in question. (A similar argument can be advanced in the domain of key analysis. Melodies may have accidentals, or notes which are not in the set defined by the key signature. Therefore, any algorithm for deciding key must place great reliance on the early notes of the tune, and not be thrown by later accidentals.) These observations can be summed up in the following

general principle which will be called the 'principle of consistency'. It is, that:

No event inconsistent with either key or metre will occur in a piece until sufficient framework (of key or time signature) has been established for it to be obvious that it *is* inconsistent.

It is hard to see how music could be otherwise. In music where such frames of reference as key and metre are used, they must be established before deviations can be effective, unless it is the composers intention to mislead. Listeners can therefore assume that early events are consistent and take them as reliable guides to both key and metre. All of the programs discussed here work according to this principle, which is to be contrasted with an alternative approach taken by, for example, Simon.

Simon's (1968) programs were also intended to carry out analysis of key and metre, and Simon and Sumner (1968) pointed out the importance of duration, repetition, and variation in identifying the metrical structure of music. Simon's approach stemmed from earlier work by himself and others (Simon and Kotovsky 1963; Restle 1967) on encoding in sequential concept attainment, reviewed together with later work along the same lines (Leeuwenburg 1969; Restle 1970; Vitz and Todd 1969) by Simon (1972). The sequential concept learning studies attempted to explain the kind of learning elicited by letter-series completion tests, where subjects are asked to extrapolate a letter sequence. In order to do this, they must infer the rule by which it has been produced. Simon's theory involved a formal language for representing such sequence-producing rules. The language was used in programs for inducing descriptions from the series, and for performing the task of extrapolation. The complexity of the descriptions involved was used to predict the difficulty which people would experience in carrying out the task. The symbolic descriptions of the series involved ideas of periodicity which transfer rather directly to the musical domain, and, in particular, to the problem of metrical analysis. In the program LISTENER, Simon (1968) made use of such rules as that long notes tend to be accented, and hence to fall at the beginning of metrical units, to infer similar sorts of descriptions of the metrical structure of a melody to those involved in the model of serial concept learning.

A major limitation on this approach is that, while in the series-completion task it is of the essence that *all* of the sequence may be assumed to conform to the rule that is to be inferred, in the case of a piece of music such an assumption cannot be made, as has been seen. Nor will a process of statistical weighing of evidence get round this problem, since a piece may actually change its metre or key, as when it includes triplets, or undergoes a modulation. An analyser that treats all the notes of the piece at once will suggest that such pieces are ambiguous, rather than revealing that the frames are different at different times.

The principle of consistency is invoked to deal with this problem, and to explain how such evidence as the occurrence of accents can be used to infer the metre of a piece, in the face of the fact that some of this evidence may be inconsistent. In saying that early evidence is always more to be relied on than late, the principle explicitly invokes the order in which the notes of the tune are heard by the listener. A computer program, analogous to a parser, working through the notes of a melody from left to right, is a uniquely suitable formalism for embodying this principle in a model of human performance.

### 3 Long notes and metre

The metre of the subject in figure 1a seems to be made so obvious by the predominance in the early notes of a rhythmic figure consisting of a long note, followed by shorter notes, followed in turn by a longer note (underlined in figure 1a). This figure suggests

a binary grouping of eighth-notes, indicated by the dotted bar lines. (Bach's own bar lines are the solid ones.) Since it is the earliest evidence for grouping above the eighth-note level, it must be *reliable* evidence as to the metre, by the principle of consistency. The comparatively low level of metric structure that results is enough to show that the long note (marked with a star) does not begin on the first beat of a metric unit, and is therefore a syncopation. A similar occurrence of isolated long notes early in the subject of fugue 15 of book I (figure 1b) seems to establish the triple grouping indicated by dotted bar lines, in time to make it clear to the listener that the late long note (starred) is syncopated.

In the earlier study Longuet-Higgins and Steedman (1971) described a model of metrical analysis, in the form of a computer program which attempts to capture these and other intuitions concerning the effect of purely durational cues on the perception of metre. The program was applied to the analysis of the forty-eight fugue subjects of Bach's *Well-Tempered Clavier*. (These were chosen as constituting a suitably varied set of short unaccompanied melodies, of impeccable musical well-formedness.) The input to this and other programs described here, rather than being the sound of a performance, was a representation of a melody as a list of notes, each note being represented by some numbers identifying the piano key used to play it, and its duration in a totally deadpan, unexpressive performance. The program gave as output a representation of the metre of the piece, in sufficient detail to determine the relevant aspects of the score, such as the position of bar lines, ties, and so on. Its results are set out in table 1.

**Table 1.** The results of the preliminary program, using rules based solely on note durations. The table shows the result that would correspond to Bach's own notation, and the result actually obtained for each fugue subject: > means that whole bars are grouped together; = means that notes are all equal in length; \* indicates an erroneous result; IAR means that the isolated accent rule applied.

Fugue	Book I			Book II		
	correct	found	comment	correct	found	comment
1	2 <sup>3</sup> /8	2/8		2 <sup>3</sup> /16	2 <sup>2</sup> /16	
2	2 <sup>4</sup> /16	2 <sup>3</sup> /16	IAR	2 <sup>3</sup> /8	2/8	
3	2 <sup>3</sup> /8	2/8		2 <sup>3</sup> /8	1/8	=
4	1/1	2/1	>	2 <sup>2</sup> .3/16	1/16	=
5	2 <sup>5</sup> /32	2.4/32	IAR	2 <sup>3</sup> /8	2/8	
6	3.2/8	2/8		2 <sup>3</sup> .3/16	2*.3/16	*
7	2 <sup>4</sup> /16	2 <sup>2</sup> /16		1/1	1/1	
8	2 <sup>2</sup> /4	1/4		2 <sup>3</sup> /8	2*.2/8	*
9	2 <sup>3</sup> /8	2/8		2/1	1/1	
10	3.2 <sup>2</sup> /16	1/16	=	2 <sup>2</sup> .3/8	2.3/8	IAR
11	3/8	3/8	IAR	2.3/16	3.2/16	(*), IAR
12	2 <sup>2</sup> /4	2/4		2 <sup>2</sup> /8	2/8	
13	2 <sup>3</sup> /8	2/8		2 <sup>4</sup> /16	4.2.4/16	>, IAR
14	2.3/4	3/4		2 <sup>3</sup> /8	2.4*/8	*, IAR
15	2.3/8	3/8	IAR	3.2/16	1/16	=
16	2 <sup>3</sup> /8	2/8		3/4	1/4	
17	2 <sup>3</sup> /8	2/8		2 <sup>3</sup> /8	2/8	
18	2 <sup>2</sup> /4	1/4		2.3/8	1/8	=
19	3 <sup>2</sup> /8	1/8	=	2 <sup>4</sup> /16	2*.2/16	*, IAR
20	2 <sup>4</sup> /16	2 <sup>2</sup> /16		2 <sup>2</sup> /4	1/4	
21	3.2/8	1/8		3.2/8	1/8	=
22	2/2	1/2		3/2	3*/2	*, IAR
23	2 <sup>3</sup> /8	2 <sup>2</sup> /8		2/2	1/2	
24	2 <sup>3</sup> /8	4/8		3/8	1/8	

The conventions used in this and subsequent tables require a brief comment. A traditional time signature is written as a fraction, for example, 6/8. The denominator of the fraction identifies the duration of the lowest metrical unit; 8, in this case, identifies the lowest unit as an eighth-note, 4 would mean a quarter-note, and so on. The numerator of the fraction indicates the number of these lowest units that make up a bar. However, the traditional time signature is rather unexplicit in indicating the subdivision of the bar. For example, the time signature written 6/8 is understood by musicians to mean, not only that the bar contains six eighth-notes, but that it is subdivided into two groups of three eighth-notes (as opposed to three groups of two). Accordingly, a related but slightly more explicit convention, in which the numerator is expressed as an ordered series of factors, has been used in table 1. So, for example, the time signature traditionally written as 6/8 appears as 2.3/8, in contrast to 3.2/8. Where the program has produced a wrong grouping, the factor which is in error is marked with an asterisk. For example, 2\*.3/8 indicates an analysis in which triplets of eighth-notes were incorrectly grouped in pairs, whereas 2.3\*/8 indicates one in which the triplets themselves were wrong. When a factor occurs repeatedly, as for example in 2.2.2/8, it is written as a power of the factor, in this case as 2<sup>3</sup>/8, for the sake of brevity.

It should also be pointed out that the program does not identify particular levels in the metric structure as those of bar and beat. These are to a certain extent arbitrary in traditional notation, and the program just produces as long a metric grouping as possible. On occasion this takes it to the level of double bars, and this is indicated in the table.

The analyses, summarised in the table, were generally of a rather low level, often only grouping the very shortest notes into pairs or triplets. In particular, the program was helpless when faced with the many fugues whose subjects consist entirely of notes of the same duration. Many of these have very clear evidence as to their metric structure, even in a deadpan performance, because of the repetition of some fragment of the melody. In the remainder of the paper, an account is developed of

Figure 1. (a) Fugue 2, book I. The metre is apparent from the relative note durations alone, particularly the 'dactyl' figure (underlined). Thus the late long note (starred) can be perceived as syncopated. (b) Fugue 15, book I. The 'isolated accent rule' operates on the arrowed notes to make the triplet metre apparent well before the arrival of the long note (starred) which would otherwise suggest a binary grouping. (c) Fugue 4, book II. The metre is apparent from the repeated occurrence of the melodic figure marked with chevrons.

melodic repetitions, and the way people use them to infer metre. The model takes the form of a second computer program in which the program already described is used in a 'first pass'.

## 4 Melodic repetition and metre

### 4.1 *Simple repetition*

An example of melodic repetition is to be found in fugue 4, book II (see figure 1c). All the notes have the same duration, so the first program draws no inference. Nevertheless, there are clues to the metre in the repetition of the melodic pattern consisting of the first three notes. (Once again, those who find the notation opaque can ignore the details, and just regard the music as a graph of pitch against time.) The repeating pattern is marked with a chevron. Such a melodic pattern will be called a 'figure', and its recurrence in a repetition will be called a 'repeat'. The first repeat seems to tell the listener that the notes are grouped in threes. The second seems to show that these triplets are in turn grouped in fours. These suggestions coincide exactly with Bach's own bar lines and time signature as shown in the figure.

Although the way that the repetition in the last example influences the perception of metre is intuitively so obvious, it raises several important points for a more general and formal account. In particular, the first repetition shows two important ways in which a repetition can involve a variation on the original figure, rather than be an identical replica. First, the repeat is at a different absolute pitch, shown by its vertical position on the musical staff in figure 1c. Second, the interval between the notes of the figure is not quite the same as that between the notes of the repeat (although the notation does not transparently reflect this fact). The interval in the figure is *one* keyboard semitone, whereas that in the repeat is *two* semitones. However, as the notation makes clear, these intervals are all *one step* in the scale identified by the key signature. A repetition of this kind is termed a 'simple' repetition, and is defined as follows:

A pair of sequences of three or more successive notes of a melody constitutes a simple repetition if all the notes before the last one are equal in duration, and if the corresponding intervals between the notes in the two sequences involve the same number of steps in the scale identified by the key signature, in the same direction.

Since an interval of a given number of keyboard semitones can be expressed as a number of scale steps only with respect to a known key signature, it follows that rules for deciding metre on the basis of this definition presuppose the identification of key, the other of the two problems tackled in the earlier work. Fortunately, the rules for key decision advanced there do not depend on any knowledge of metric structure. There is therefore no danger of circularity in the rules. Once the key of a piece is known, there are a number of simple ways of deciding the size in scale steps of the interval between any two keyboard pitches.

Having observed that notes 4, 5, and 6 in figure 1c constitute a repetition of notes 1, 2, and 3, the obvious conclusion that the eighth-notes are grouped in threes, with the accent on the first and fourth notes, rests on the assumption that *the metric accent will fall at the same point in both figure and repeat*. This allows the inference that the metre is triple. Furthermore, there is no temptation in this example to hear the accent on anything but the *first* note of the figure and its repeat, so the following rule suggests itself:

The metric grouping to be inferred from a repetition has a period equal to the separation of figure and repeat, and places the accent on their first notes.

Although some qualifications will have to be made, particularly to widen the definition of what can constitute the beginning of a repetition, this is the basic rule by which the program works.

Another basic feature of the program is illustrated by the second repetition in this subject. The initial figure is repeated again by notes 13, 14, and 15. By the above rule, it can be concluded that the triplets established by the earlier repetition are, in their turn, grouped in fours. The accents of the quadruple groups occur on notes 1 and 13. This takes the analysis quite correctly to the level of Bach's own bars. However, notes 13, 14, and 15, also constitute a repetition of a figure containing notes 4, 5, and 6. It would be wrong to infer a metric unit on this basis. A repetition is therefore taken to involve the *earliest* possible figure in such cases.

There are other repetitions among the notes of this subject which must *not* be taken as evidence for metric structure. Notes, 7, 8, and 9 of figure 1c could be seen as repeating notes 2, 3, and 4. However, this is ruled out of court on two counts. First, the repetition of the first three notes by the second three happens first, and by the principle of consistency it cannot be overruled by later evidence. Second, metric groupings with prime factors other than two and three are held to be unreasonable and are not accepted by these rules. (Although, in other music than Bach's, groupings of five, seven, and even higher prime factors are allowed, these are generally at a very high level and so this does not constitute a serious limitation on the theory.) The repetition in question here would propose an unreasonable grouping of eighth-notes into fives, so it is ignored. Another potential repetition of notes 2 to 6 in notes 11 to 15 must also be ignored, since it conflicts in phase with the previously established triplets.

#### 4.2 *An algorithm*

It is possible to embody all of these observations in an extremely simple algorithm for metrical analysis, which will form the basis of the program described in the next section. Typically, the algorithm must deal with the following situation. Some part of the melody will have been dealt with already, and a partial knowledge of its metric structure will have been gained. The remainder of the melody must next be examined in order to decide whether the sequence with which it begins is a repetition of anything that has been encountered previously. If a repetition is found, then it may establish a new and more complete metre. If not, the algorithm can pass on with an unchanged metric unit. In either case, only those sequences which begin on an accent of the established metre need be examined, since evidence from all other sequences is excluded by the principle of consistency.

These remarks can be summarized as four characteristics of the algorithm:

- 1 The algorithm should consider the remainder of the tune in 'bites' of whatever metre has been established, and ask whether the sequence of notes starting at the beginning of such units constitutes a repetition of anything previously encountered.
- 2 In searching for a corresponding figure in the earlier melody, only the sequences which start at the beginning of a metric unit need be considered.
- 3 Only those figures in the earlier part of the melody which are separated from their repeat by a 'reasonable' number of units need be considered.
- 4 Because of the remarks concerning the primacy of early figures over later ones, it is clear that the search of the earlier part of the melody should be carried out from earliest to latest.

In the subject under discussion (figure 1c), the above algorithm works as follows. Nothing much happens until the fourth note, when a repetition of the first three notes is detected, and so a metre of three times the basis sixteenth-note metric unit is established. The analysis therefore continues in 'bites' of this new unit, corresponding

to the dotted bar lines in the figure. No repetition is found to begin on the first note of the third such unit, and sequences beginning on its second and third notes are never even considered. The same goes for the fourth triplet unit. When the fifth unit is compared with earlier ones, in order from earliest to latest, it is immediately found to constitute a repetition of the first unit. A metre which groups the previously established triplets in fours, represented by the time signature 4.3/16, is accordingly established and the analysis proceeds to the end of the tune in terms of this unit.

When this algorithm is applied to other subjects among the forty-eight, it soon becomes clear that the definition of repetition that has been given so far is too restricted. In particular, as well as the exact or 'simple' repetition of a sequence of scale intervals, repetitions involving certain kinds of variation on the figure must be considered.

#### 4.3 Variant repetition

The subject of fugue 20 of book I (see figure 2a), involves a more general kind of repetition than those that have been discussed so far. The first important clue as to the metre of this piece is a purely durational one, of the kind recognised by the old program. There is a 'dactyl' including notes 5, 6, and 7, which sets up a perfectly correct binary grouping of the eighth-notes, shown by the dotted bar lines. (Bach's own bar lines are the solid ones.) This is enough to show that the piece begins with an eighth-note rest—in other words, the first note sounded falls on an unaccented beat. But this is as far as the old rules take the analysis.

At note 8, there is a repetition of notes 3, 4, and 5. (The durations of notes 5 and 10 are not the same, but since they are the last notes of the sequences, and it is only intervals *between* successive notes whose size and duration matter, they still take part in the repetition.) If the algorithm is allowed to construct a metric grouping of pairs of the previously established units, then it will produce an incorrect metre. The problem is that, while one does seem to hear this repetition, one hears it as starting

Figure 2 consists of three musical staves labeled (a), (b), and (c).  
 (a) Fugue 20, book I: Bass clef, 4/4 time signature. Notes are numbered 1 to 10. A solid bracket groups notes 3, 4, and 5. A dotted bracket groups notes 8, 9, and 10. Vertical dotted lines mark the boundaries of these groups.  
 (b) Fugue 24, book I: Treble clef, 4/4 time signature. Notes are numbered 1 to 10. A dotted bracket groups notes 3, 4, and 5. A solid bracket groups notes 8, 9, and 10. Vertical dotted lines mark the boundaries of these groups.  
 (c) Fugue 17, book II: Treble clef, 4/4 time signature. Notes are numbered 1 to 12. A solid bracket groups notes 3, 4, and 5. Another solid bracket groups notes 8, 9, and 10. Vertical dotted lines mark the boundaries of these groups.

Figure 2. (a) Fugue 20, book I. The simple repetition (solid bracket) would give half-bars with the wrong phase. The groups preceding the figure and repeat are rhythmically identical and can be included (dotted brackets) as part of a variant repetition, with the correct phase. (b) Fugue 24, book I. The corresponding intervals between the notes of the variant part (dotted brackets) must be equal in duration and direction. A variant repetition must end in a simple repetition (solid brackets). (c) Fugue 17, book II. No variant part can be included in this simple repetition (solid brackets), because the corresponding intervals would be in opposite directions.



somewhat earlier. In fact, notes 5, 6, and 7 seem to repeat the very beginning of the piece. However, the figure and the repeat are no longer the same sequence of intervals. The first note of the figure is a rest, unlike the first note of the repeat, although the remainder of the sequence of intervals is identical. It is therefore necessary to allow the possibility that a repetition constitute a *variation* upon the original figure, particularly at its beginning. Such a variation is termed a ‘variant’ repetition, in contrast to the earlier ‘simple’ kind.

A variant repetition can ring the changes on a figure in more subtle ways than mere replacement of notes by rests, as fugue 24 of book I shows (see figure 2b). (Again, the solid bar lines are Bach’s own.) Although the subject contains a long note consistent with the metre, it comes rather late in the melody, and is played as a trill. The isolated semitone intervals are not by themselves evidence for repetition, since a repetition must contain at least *two* intervals between *three* notes. But the sequence starting at note 10 is a simple repetition of the figure beginning on note 6. (The figure and repeat actually overlap, and this is allowed, although it does not add anything to the metrical analysis.) The simple repetition alone would cause an incorrect grouping of the notes, with the wrong phase. Again, as in the last example, the figure and repeat are actually heard as starting rather earlier, at notes 4 and 8, respectively. In fact, the intervals between notes 8, 9, and 10 constitute a variation upon those between notes 4, 5, and 6. Together with the simple repetition, which follows immediately, they constitute a variant repetition, resulting in a metric grouping of period four, with the correct phase.

There are strong constraints upon the note sequences that may count towards variant repetition. Most importantly, every variant repetition must finish in a simple repetition, as this one does. Furthermore, notes may only be included at the beginning if they are: (a) of the same duration in both figure and repeat, and (b) if the interval between them is in the same direction—either ascending or descending. This last clause is extremely important, and is the reason that the variant repetition in the above example cannot be taken to include notes 3 and 7: the intervals between notes 3 and 4, and 7 and 8, are in opposite directions. A similar case occurs in fugue 17 of book II (see figure 2c) where the preliminary rules, based on relative duration alone, establish the indicated binary grouping of eighth-notes (because of the dactyl at note 4). The simple repetition of notes 4 to 7 by notes 9 to 12 is correctly prevented from including the preceding pairs of eighth-notes as a variant part: the intervals involving these notes would be in opposite directions in figure and repeat.

The above definition of variant repetition can be summed up in the following rule:

Two sequences of notes in a melody constitute a variant repetition if the corresponding notes (except possibly the last) are equal in duration, and the corresponding intervals are the same in direction, and the sequences end with a simple repetition.

The algorithm described before applies, unchanged, with this wider definition of repetition.

#### 4.4 Scales and alternation

There are two kinds of note sequence which must be treated as special cases by the algorithm. The first of these is the scale. A sequence of notes of uniform duration forming a scale will always contain repetitions of *all* periods, if it is long enough, as illustrated in figure 3a. The algorithm described so far will always infer a binary metre from such a scale (since the first significant repetition encountered will be of period two). This can be too hasty a conclusion, as fugue 6, book II shows (see figure 3b). It involves an early scale starting on note 6, despite being in triple time.

There is no temptation to make any metrical inference at all on the basis of the scale, and judgement is deferred until it is realised that the figure starting on the first note is repeated at note 7. At this point it becomes clear that the notes are grouped in sixes. Moreover, *all* putative repetitions which begin in the middle of a scale, as opposed to beginning at one of its ends, are suspect. For example, in fugue 18, book II (see figure 3c), the apparent repetition of notes 8, 9, and 10, by notes 11, 12, and 13 would lead to a grouping of the wrong phase. Nor is it a repetition that one hears, so it must be excluded by the rules. Nevertheless, in both these examples, the repetitions that one does hear *do* begin in the middle of scale passages (see figure 3). The difference seems to be that they include substantial simple repetitions *besides* the scale fragment with which they begin, whereas the spurious ones do not. It appears, therefore, that a figure or a repeat *may* begin in the middle of a scale, but that if it does, the intervals involved in that scale count only as a variant repetition. Unless there is a substantial simple repetition later on, no conclusion may be drawn. The consequence for the algorithm is that, as the melody is examined in units according to the metre established so far, special care must be taken if the unit in question starts in mid-scale. When it does, it will be necessary to wait and see if there is a substantial further repetition at the end of the scale. The algorithm seems to correspond closely in this respect to the way one hears this piece: the repetition of the first six notes of the last example starting at note 13 does indeed seem only to be established in retrospect, when the simple repetition arrives, starting at note 15.

Figure 15 of book II suggests another variety of sequence that must also receive special treatment, as a whole (see figure 4a). None of the rules of the preliminary algorithm, based solely on duration, apply to this subject, but notes 1, 2, and 3 are apparently repeated by notes 5, 6, and 7. This gives a grouping of the notes into fours, with a totally absurd phase. It is absurd because every other note in the first two bars is the *same* pitch, namely D, and the grouping in question accents the repeating same pitch. In such 'alternating' sequences the repeated pitch generally falls on the *unaccented* beat of a duplet rhythm. The psychological reason for this is obvious if one tries to hum the tune with the repeating pitch accented—it sounds odd to place such emphasis on notes which convey so little new information.

Figure 3 consists of three musical examples, (a), (b), and (c), each showing a scale on a treble clef staff with a key signature of one sharp (F#).  
 (a) A scale of 12 notes. Below the staff, a bracket underlines the first six notes (1-6). A second bracket underlines the last six notes (7-12). A third bracket underlines the last three notes (10-12).  
 (b) A scale of 12 notes. Below the staff, a bracket underlines the first six notes (1-6). A second bracket underlines the last six notes (7-12). A third bracket underlines the last three notes (10-12).  
 (c) A scale of 15 notes. Below the staff, a bracket underlines the first six notes (1-6). A second bracket underlines the last six notes (7-12). A third bracket underlines the last three notes (13-15).

Figure 3. (a) The multiple repetitions within a scale (solid brackets) are ignored. (b) Fugue 6, book II, and (c) fugue 18, book II. A repetition (brackets) may begin in the middle of a scale (underlined), but the intervals of the scale only count as variant repetition.

The rules therefore treat alternating sequences as a special case, similar to scale sequences. The rule is that such sequences indicate a binary grouping of the notes, with the accent falling on the nonrepeating pitches. It turns out that this rule must only apply to alternating sequences which include at least three successive occurrences of the repeating pitch separated by notes both of which are higher, or lower, in pitch. Such sequences as the one underlined in figure 4b, in fugue 18 of book II are not heard as 'alternating' in this way, whereas the alternation at the beginning of fugue 15 of book II *does* include such a sequence, between notes 6 and 10.

The same subject raises a further problem. The above rule groups the notes in pairs as shown, with the repeating D unaccented and the initial rest established. Once this has been done, there is an incontestable variant repetition of a four-note figure beginning on the initial rest, by notes 4 to 7. If this is accepted then it will cause the algorithm to set up a metric group containing *four* pairs of notes, whereas Bach's time signature indicates that the pairs are grouped in *threes*. This cannot be held to be a failure of the rules, for one does seem to hear the repetition quite clearly, and draw the inference that the metre is quadruple at this point in the tune. However, one *also* hears the repetition of notes 12 to 17 by notes 18 to 23, which suggests the triple grouping that Bach indicates, and for once the later evidence is allowed to override the earlier. It seems that another way in which alternating sequences are special is that any repetitions that they include are only evidence for their *internal* metric structure: no conclusions as to the structure of the piece as a whole can be drawn, until the sequence has finished. They are similar to runs in this respect. With this qualification, the first evidence that the algorithm takes into account is the variant repetition of the first four notes (including the initial rest) by notes 12 to 15. This gives a bar of *twelve* sixteenth-notes—twice the size of Bach's own bar, but with entirely correct phase. The algorithm should at this point go back and examine the internal structure of the two large bars that it has produced. It should infer that the metric structure of the first twelve notes involves three groups, each of two groups, which in turn contain two notes, whereas the second twelve-note bar contains *two* main groups, each containing three pairs of sixteenth-notes. (This structure is illustrated beneath figure 4a.) However, there are difficulties associated with going back over large metric units and filling in 'missing levels' of metric structure, and this refinement of the algorithm has not been included in the program.

Figure 4. (a) Fugue 15, book II. The 'alternation sequence' (underlined) sets up a binary grouping (dotted bar lines), with the accent on the nonrepeating pitch. Repetitions within alternation sequences are ignored, so the first evidence for a higher metre is the variant repetition (bracketed), which yields the metre 6.2/8. The detailed metric structure of the melody is indicated below the staff. (b) Fugue 18, book II. The underlined sequence does *not* constitute an alternation sequence.

## 5 Program and results

The algorithm has been implemented in POP-2, a high-level list-processing language, as a simple computer program. It takes as input a list representing the simple duration, and the piano key used, for each note of a melody, and gives as output an indication of the metric structure of the piece, analogous to information implicit in the time signature, bar lines, beams, and ties, of a traditional musical score. The program uses the earlier rules in a first 'pass', and uses the earlier key-identification program to interpret the keyboard intervals between the notes of the input as number of scale steps in whatever key the melody happens to be. The new program has been run on the same forty-eight fugue subjects of Bach's *Well-Tempered Clavier* as the earlier programs.

The rules yield a higher level of metric analysis in the second pass for twenty-one of the forty-eight fugue subjects. Among these the rules produce a correct metric analysis in sixteen cases. In five cases—fugues 3, 19, and 21, of book I, and 16 and 21 of book II (see figure 5)—the new rules produce wrong analyses. These results are summarised in table 2, with the use of the same conventions as in table 1.

The five errors raise interesting problems for future work. Fugues 16, book II, and 19, book I, yield wrong analyses which, it could be argued, human listeners themselves are drawn into, and where it seems to be the composer's intention to mislead. In fugue 16, book II, this has the effect of establishing a metre of the correct triple period, but of incorrect phase, for there is no way for the rules to infer the initial rest, shown in the score (see figure 5a). The human interpreter may also make this

**Table 2.** The results of the program in which repetition was used. The result which would correspond to Bach's notation, the result obtained in the first pass, and the results obtained in the second pass are shown for those subjects for which the program gave new additional metric information: > means that whole bars are grouped together; \* indicates an erroneous result; AR means that the alternation rule applied.

Fugue	Correct	Previous	New	Remarks
Book I				
2	2 <sup>4</sup> /16	2 <sup>3</sup> /16	2 <sup>4</sup> /16	
3	2 <sup>3</sup> /8	2/8	2*.2/8	*
5	2 <sup>5</sup> /32	2.4/32	2 <sup>2</sup> .4/32	
10	12/16	1/16	2/16	AR
15	2.3/8	3/8	3.2.3/8	>
19	3 <sup>2</sup> /8	1/8	2*/8	*
20	2 <sup>4</sup> /16	2 <sup>2</sup> /16	2 <sup>3</sup> /16	
21	3.2/8	1/8	6*/8	*
24	2 <sup>3</sup> /8	4/8	4.4/8	>
Book II				
1	2 <sup>3</sup> /16	2 <sup>2</sup> /16	4.2 <sup>2</sup> /16	>
4	2 <sup>2</sup> .3/16	1/16	4.3/16	>
7	1/1	1/1	2/1	>
10	2 <sup>2</sup> .3/8	2.3/8	2 <sup>2</sup> .3/8	
11	2.3/16	3.2/16	2.3.2/16	>
12	2 <sup>2</sup> /8	2/8	2 <sup>2</sup> /8	
15	3.2/16	1/16	6.2/16	>, AR
16	3/4	1/4	2.3*/4	*, >
17	2 <sup>3</sup> /8	2/8	2 <sup>2</sup> /8	
18	2.3/8	1/8	12/8	>
21	3.2/8	1/8	6*/8	*
24	3/8	1/8	3/8	

mistake, in the absence of some expressive guidance from the performer in the form of stress on the metrically accented notes. A similar case is that of fugue 19, book I (figure 5b). The sequence which immediately follows the three eighth-note rests is repeated by the sequence from note 7. These sequences form part of a variant repetition which results in a binary grouping, at odds with Bach's time signature of 9/8. This is because the passage following the rests (underlined) has the form of two parallel scales, interleaved with one another. Such sequences are common in Bach's music and generally have the binary structure that the repetition rules will always ascribe to them. In this case, Bach simply breaks this rule, and the listener receives a great surprise when the 'answer', or second statement of the melody, joins in after a period of *nine* beats, making it clear that the initial impression must have been quite wrong. Again, this is an instance where the rules point out a case where the performance must give some help, say by playing the metrically accented notes louder than the rest, if the listener is not to be misled.

Figure 21 of book II (figure 5c), is a more ambiguous case. The repetition has the correct period but the wrong phase. The true metre seems to become obvious in the third bar, when one realises that the first of each of the same-note pairs must fall on an *unaccented* beat. However, it is not clear how to include this intuition in the rules.

Figure 3 of book I (figure 5d), illustrates a serious flaw in the algorithm as described, and suggests a remedy. The early dactyl—notes 8, 9, and 10—establishes the correct pairing of eighth-notes indicated by the dotted bar lines. The descending

Figure 5. Errors made by the program: (a) fugue 16, book II; (b) fugue 19, book I; (c) fugue 21, book II; (d) fugue 3, book I; (e) fugue 21, book I.

parallel scale sequence which immediately follows is just long enough for a *second* level of binary grouping to be inferred on the basis of the repetition of notes 11 to 13 by notes 15 to 17. This binary grouping has the wrong phase, and is *not* a mistake that a human listener could possibly make.

The error is exactly the same as the one that was avoided by prohibiting repetition within simple scales. The sequence here, from note 11 to 17, is also a sort of scale. Although it is perfectly justifiable to infer a metre from the repetition of the low-level units, repeating at a remove of one scale step, no *further* inference can be drawn within the sequence, any more than it can within a simple scale. A similar occurrence, in fugue 11, book II results in a bar twice as long as Bach's own. Another case in point is fugue 16, of the same book (figure 5a). Besides the triplet metre of correct period but incorrect phase, which is set up by the repeating units, the fact that they form a scale sequence causes a second level of binary grouping. Again, this gives a bar twice as long as Bach's own. Neither of these last two analyses is particularly wrong but they are not really justified. The rules should be expanded to recognise such 'concealed scales', and inhibit inferences within these sequences, in a manner analogous to that implemented for simple scales. However, these last two examples show that the definition of what counts as a complex scale is going to be very elaborate, and no such extension of the rules has been implemented.

Fugue 21 of book I (figure 5e) gives rise to an error which shows that definition of melodic repetition set out in this paper is, not surprisingly, too narrow. The perfectly correct simple repetition found by the program is not part of a variant repetition as defined, and the resulting metre has the wrong phase. However, the error is due to the limitation in the definition of a repetition to interval sequences involving the same durations. The second bar does in some sense repeat the first bar, but it involves a change in the number of notes, and so these rules will not find it.

A further limitation of the program is that it applies its two kinds of rules in two separate passes through the melody. While it is believed that all of the rules described here are basically independent of this aspect of the program, and a one-pass algorithm has been devised, the corresponding program is still being developed.

## 6 Conclusion

This program is intended to constitute a psychological theory of our perception of metre in melody. Its results can be regarded as a detailed prediction of human performance on the task of identifying the metre of these and other melodies from tonal performances in which explicit cues (such as variations in loudness and timing) are lacking. The prediction is that where the program does not infer the metre that the score indicates, then either the melody is ambiguous, or the composer has exercised the artist's privilege to break the rules. In either case, the melodies are simply misleading in the absence of some expressive guidance from the performer, and people will make the same mistakes. Conversely, where the program succeeds, the prediction is that people will hear the appropriate metre, even in the absence of such expressive cues. We are currently investigating people's performance on the task, using artificially constructed uninflected performances of the melodies.

To summarise: the program takes account of certain simple constituent structures of melodies, namely scales and alternations. (Certain of the errors indicate that its repertoire of such constituents must be expanded to include complex or concealed scales.) In the terms of the analogy with language with which the paper began, such constituents may be viewed as 'syntactic' elements in a 'grammar' of music, and the program can be viewed as parsing a melody according to this syntax, in order to determine a simple aspect of its meaning, namely the metre. As a parser it makes no separation between parsing and interpretation: both are done together as early in the

melody as possible. This aspect of the program is to be contrasted with the approach taken in some other studies. It is expressed as a 'principle of consistency', which explains how early evidence in the melody can be used to set up a representation of the metre, against which later syncopations can be understood. The principle may be of general relevance both in studies of sequential concept attainment and in the study of natural language. In English, the very first word of a sentence is often enough to identify it as a question, statement, or command. As Winograd (1972) points out, this fact can be taken advantage of immediately, and is rarely gainsaid by later evidence.

**Acknowledgements.** This work was done in the Theoretical Psychology Unit of the School of Artificial Intelligence at the University of Edinburgh. It was supported by a grant from the Royal Society, and used the computing facilities of the School of Artificial Intelligence, which were provided by the Science Research Council. It has been influenced at every stage by conversations with Christopher Longuet-Higgins, who supervised the PhD thesis (Steedman 1973) of which it formed a part. Thanks also to Stephen Isard and Phil Johnson-Laird for all kinds of musical exchanges.

### References

- Leeuwenberg E L L, 1969 "Quantitative specification of information in sequential patterns" *Psychological Review* 76 216-220
- Longuet-Higgins H C, 1976 "The perception of melodies" *Nature (London)* 263 646-653
- Longuet-Higgins H C, Steedman M J, 1971 "On interpreting Bach" in *Machine Intelligence 6* Eds B Meltzer, D Michie (Edinburgh University Press, Edinburgh)
- Restle F, 1967 "Grammatical analysis of the prediction of binary events" *Journal of Verbal Learning and Verbal Behaviour* 6 17-25
- Restle F, 1970 "Theory of serial pattern learning" *Psychological Review* 77 481-495
- Simon H A, 1968 *Perception du Pattern Musical par "AUDITEUR"* *Sciences de l'Art Tome V.2* 28-34
- Simon H A, 1972 "Complexity and the representation of patterned sequences of symbols" *Psychological Review* 79 369-382
- Simon H A, Kotovosky K, 1963 "Human acquisition of concepts for sequential patterns" *Psychological Review* 70 534-546
- Simon H A, Sumner R K, 1968 "Pattern in music" in *Formal Representation of Human Judgement* Ed. B Kleinmuntz (New York: John Wiley)
- Steedman M J, 1973 *The Formal Description of Musical Perception* unpublished PhD thesis, University of Edinburgh
- Vitz P C, Todd R C, 1969 "A coded element model of the perceptual processing of sequential stimuli" *Psychological Review* 76 433-449
- Winograd T, 1972 *Understanding Natural Language* (New York: Academic Press)