Musical Expression in Automated Composition of Phrases

Brian Thomas McLintock
John W. Roach

TR 87-34

Musical Expression in Automated Composition of Phrases

by

B. T. McLintock
J. W. Roach

Virginia Tech
Department of Computer Science
TR-87-34
November 1987

ABSTRACT

Music composed by computers has always been lacking in "musical" qualities: mood, emotional expression and a sense of purposefulness or goal. A musical expert system, called EMOTER, is the first attempt to address these important musical aspects. EMOTER receives as input a list of moods (e.g., happy, lively) and generates melodic passages intended to evoke those moods in an organized, coherent fashion. EMOTER composes the basic units of music called phrases.

The program uses the mood-specification from a theory due to Deryck Cooke to derive a few motifs (very primitive melodic material) exemplifying the moods and computes a number of musical attributes to guide its compositional choices. A theory of emotion due to Leonard Meyer further helps plan the phrase. The theory states that an emotional response is stimulated in a listener when expectations about the progression of the music are first established and then inhibited (with the understanding that the expectations will eventually be fulfilled). A melodic passage is composed using the selected motifs, attributes and emotional theory to create a "skeletal" phrase. This is embellished and developed (also using the attributes and theory) to flesh-out the bare melodic material into a passage that embodies the musical characteristics of the mood-specification.

Results with EMOTER are excellent. Many musical phrases comparable to music of normal composers are generated from a single mood-specification. More theory is needed, however, before the full complexities of human-composed music are sufficiently captured in code for EMOTER to pass a Turing test in music composition.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Problem

This thesis initiates an orderly investigation into the fuzzy, mysterious realm of aesthetics and emotional expression from the viewpoint of computer science. A program called EMOTER tests the theories put forth, using music as a vehicle for the tests. It takes as input a list of moods and produces as output phrases (in the common-practice style of 18th-century Western classical music — hereafter called "Classical" in the thesis) embodying a musical representation of those moods. In addition, the phrases are designed to have emotional swings, a sense of direction and destiny — in short, to sound as unmechanically composed as possible. The phrase was chosen as the melodic unit because it is

> ...the smallest structural unit ... a kind of musical molecule consisting of a number of integrated musical events, possessing a certain completeness ...

[Schoenberg 67, p. 3].

Three theories formed the basis of EMOTER: Leonard Meyer's theories of emotion and meaning, and the ideas of several authors — notably Deryck Cooke — on the ways that music suggests certain moods. The third theory was developed while working on the design of EMOTER. The problem of finding a relationship between time and pitch for the rhythmic aspect of composition forced us to invent our own theory of rhythm.

Preliminary investigation into automating the first two theories has shown the need for considerable background work. In particular, most of the time has been spent developing an environment of rules and representations to manipulate and store the moods and the music. Since the main purpose of this thesis is *not* simply to generate phrases in the Classical style, the implementation of a body of rules just to do this seemed a waste of time and effort. A minimally suitable subset of conventional music theory rules has, however, been written so that the generated music will not sound so alien that perception of the intended effects is seriously clouded. Even this relatively small body of rules has turned out to be a large undertaking.

1.2 Motivation for this Research

Artificial intelligence researchers want to model (with the aid of computers) how the mind works. Usually their interest is understood to be in the deductive or epistemological realm, but emotion is also a valid and important part of our being. With the exception of Parry ([Faught 77, Colby 75]), AI research has not concentrated on this area; no composition programs written so far claim to be expressive in the musical sense. To the best of our knowledge, this thesis represents the first serious attempt to compose music that intentionally expresses a mood and has emotional content. If the attempt proves successful then we will have shown that an accurate mechanical or logical model of human emotional response is possible.

1.3 Brief Description and Evaluation of the Results

EMOTER produces, for a typical mood specification, between five and one hundred phrases, each roughly five seconds in length (if played at the tempo specified). The program could easily produce many more, but EMOTER has been instructed to output only its "best" phrases. The example below contains a typical phrase with its mood specification.

`(outgoing protesting tense active)`

Do the phrases exhibit consistent mood throughout? Do they seem "emotional?" Are there "goals" toward which the music progresses? Just how "good" are they? These questions cannot be accurately answered with complete objectivity, for the musical experience is by its nature a subjective one. We will admit that the phrases are fairly simple and lack some of the grace of, say, Mozart's. Nevertheless, in our opinion the phrases produced by EMOTER are generally musical and express the given moods.

1.4 Research on Representations Appropriate for Music

According to the Explicit Description Hypothesis one cannot hope to understand or control a thing until one has an adequate representation of it. To be able to explain to a computer how to compose music it is therefore necessary to provide the computer with a representation for the music. Specifically, a representation of the relationship between music and emotional response is needed if composition of "musical" music is ever to be achieved. This presupposes the presence of a grammatical representation for the mechanics of music production. The development of suitable representations for the "semantic" and "syntactic" aspects of the problem — the moods and their meanings, and the music notes themselves — will be discussed briefly here, and in more detail in Chapter 3.

1.4.1 Representing Moods

The moods should be represented in such a way that useful information about them can be made available to help the program decide how to write the phrase. How is something like a feeling or mood explained to a machine? Our method uses a set of well-defined properties of music that can be manipulated and constrained by certain operations and controls. Music has a plethora of such *attributes*. Which ones are appropriate for "expressing" feelings, and how do they relate to one another? We will explain the choices we made and explore their relationships.

1.4.2 Representing Notes

Representing the music notes seemed easier than representing moods — after all, many music representations for computers exist. But few exist for an artificial intelligence language such as Lisp or PROLOG. A truly useful representation for notes should be flexible so that different representations of the same notes might be possible, expandable to allow addition of other music notation, complete enough to express any conventional music, readable by humans to make editing fairly easy, and manipulable by a computer program (so that the representation must have a consistent syntax). We devised such a representation, called NOTES, which fulfills those requirements quite well. For easy calculation of the basics (pitch and time), however, a second representation was invented.

1.5 Organization of the Thesis

Chapter 2 discusses the theories used and adapted for this thesis. Following that are chapters covering the operation of the program in detail, including representations, operations, metaknowledge and control. Chapter 8 discusses the Artificial Intelligence tools, techniques and models we used to design and build EMOTER. Chapter 4 lists the general plan of EMOTER's operation. The next three chapters discuss the theory and application of the music attributes, rhythm and developments, respectively. Chapters 4 through 7 include a detailed trace of an actual program run that generates the musical phrase shown above. Chapter 9 lists ideas for future work to extend what has been done here. Finally, there are appendices containing a glossary of terms, some typical mood specifications, a detailed description of the note representations, a considerable sampling of music composed by EMOTER and a bibliography. The glossary contains definitions specific to this thesis; we recommend consulting it whenever encountering a word of dubious meaning or found in a confusing context.

Chapter 2

Background: Theory

2.1 Introduction

This chapter provides the historical and theoretical background for the remainder of the thesis. The first section briefly surveys previous attempts at computer composition. The history of sound synthesis is not included since we are not concerned with that here. In the remainder of this chapter the aesthetic and musical theories forming the foundation of this thesis are explored. First, the factors that influence the meaning of music are introduced. The following two sections discuss the theories of Deryck Cooke [Cooke 83] and Leonard Meyer [Meyer 56], respectively. Their ideas form the basis for much of this thesis and the program that implements it. Then conventional music theory is briefly discussed and related to the theories of Meyer. Finally a method for representation of mood in music is introduced. This representation is our own, but we owe a debt to Cooke for giving us a start with his "vitalizing agents."

2.2 A Short History and Critique of Automated Composition

Historically, automatic composing procedures have been based on one or more of the following methods:

- clever use of stochastic processes [Beauchamp 69, Ch. IV], [Pinkerton 56]
- a grammar expressing original or established "music theory" or style rules [Brooks, Hopkins, Neumann et.al 57], [Zaripov 60]
- generation of a structure to organize the music [Hiller & Isaacson 79]
- generation of an accompaniment to previously-composed music [Rader 74]

While academically interesting, the above methods are in our opinion musically unsatisfying. No attempts have been made to instill the music with any "emotional content" or "mood" in a direct manner. To satisfy the curious reader (and attempt to reassure the skeptical one), we shall now summarize a few of the more influential projects in the above areas.

2.2.1 The Use of Randomness

Random choice has been the most common method of automated music composition. The earliest attempts were based on this method, but they were never intended to be taken seriously [Hiller & Isaacson 79, p. 54]. Mozart's *Musikaliches Würfelspeil*, for example, consisted of short fragments of music that would sound reasonable no matter how they were arranged. Dice rolls randomly determined how to paste them together. The first "serious" application of random choice to automated composition had to wait for Hiller and Isaacson's famous series of experiments [Hiller & Isaacson 79]. One of their techniques made use of nth-order transition probabilities. By looking back at the last n events (notes) and based on previous music containing that same string of events, a weighted random choice of the next note was made. This technique depended on the syntactic principle of past experience blindly followed to choose the sequence of notes. Hiller and Isaacson also used a technique called generate and test. A pitch was randomly chosen (the randomness in this case being merely a convenience), then subjected to a number of tests for appropriateness according to a subset of the rules of music theory (e.g., voice-leading). Pinkerton [Pinkerton 56] used information theory and examined different types of randomness and their effects on melodic continuity. However, information theory is an analytical method and Pinkerton admitted that melodies generated using his ideas were bland and lacking in organization.

2.2.2 Use of a Musical Grammar

A much more intellectually appealing method than randomness (at least from an Artificial Intelligence viewpoint) for composing music is the use of a musical grammar, especially a generative one. Music has often been likened to a language, and it was only natural that many treated it that way when approaching the task of automated composition. [Holtzman 81], [Roads 78] and others have written generative grammars for melodies, often in the style of traditional music. Grammars were felt to be ideal for imposing a structure on the music, a feature which randomly-composed music lacked. [Roads 79] also used a grammars for the representation of music. While music generated by a grammar undoubtedly can demonstrate structural integrity, it cannot claim to be expressive based on its grammar alone.

2.2.3 Algorithmic Composition

[Hiller & Isaacson 79] and [Holtzman 81] (and others [Anon. 62]) have also used a technique known as algorithmic composition. As the name implies, the programmer would write an algorithm to tell the computer exactly what to compose. Certain mathematically regular styles of music, principally avant-garde, are particularly amenable to this technique, although a few tried to copy the style of more traditional music. The reaction of audiences to this method of composition and to composition using grammars has been disappointing [Anon. 60].

2.2.4 Generation of Music Accompaniment

Programs to harmonize existing melodies or to generate a contrapuntal line to go with a previously-composed melody [Rader 74] have been written. They typically implement a set of rules about voice-leading, chord-progression and the like. The accompaniments produced were adequate but not, in our judgement, up to the contrapuntal standards of, say, Bach.

2.3 What is music about?

Music may be thought of as a combination of several high-level or abstract aspects: Mood, Emotion, Meaning, Creativity, and Aesthetics. [Cooke 83] covers mood (although he calls it "emotion") and [Meyer 56] deals with emotion and meaning. Early music theory books actually taught music aesthetics [Goetschius 02], [Rameau 71], and modern texts generally continue that tradition. Although we have some thoughts about creativity, we will not discuss them here. These five concepts will be covered first since their meanings as used here may differ from the conventional interpretation.

2.3.1 Mood

According to Meyer, *mood* is a relatively permanent and stable designation of one or more separate states of feeling. Mood establishes a set of norms in the music against which more transitory emotional events can be displayed, somewhat like the backdrop for a play. [Hevner 35] developed a "cycle" of eight sets of moods; [Cooke 83] has a small vocabulary of moods. Because music is abstract these moods cannot be given accurate names in a spoken language; nevertheless, a representative list might include the following mood "axes" or opposites:

lethargy	$\Longleftrightarrow$	liveliness
calmness	$\Longleftrightarrow$	tension
levity	$\Longleftrightarrow$	heaviness
happiness	$\Longleftrightarrow$	sadness
smoothness	$\Longleftrightarrow$	abruptness
hesitancy	$\Longleftrightarrow$	confidence

Each passage (contiguous group of notes) has one or more moods associated with it. More moods can potentially be associated with a longer passage. Each time the passage is reiterated it reinforces the moods of the original passage by familiarizing the listener with the initial pattern of notes and by the relationship of the reiteration to the original (strict repetition – static, sequence – dynamic, unpatterned change – chaotic).

Mechanically-composed music has — until now — not paid any attention to mood. Hevner's and Cooke's ideas suggest ways that mood might be mechanically generated.

2.3.2 Emotion

Unlike mood, *emotion* is temporary and undifferentiated (according to [Meyer 56]). That is, emotion changes almost constantly, but it has only one "axis." Even though emotion is a one-dimensional quantity, measuring it is not easy. Meyer mentions physical responses that are evidence of emotional responses, but he points out that the same physical manifestation can be in response to a wide emotional range (tears may accompany laughter or sorrow, for instance). Thus emotion as defined by Meyer, although undifferentiated, does have pleasant ("positive") and unpleasant ("negative") "poles."

2.3.3 Meaning

The word "meaning" in connection with music has been overused. Traditional music theory considers it to be the functional harmony and, to a lesser extent, the melodic content of the music. That is, the meaning of a piece of music is (generated by) its formal structure. While that may be true as far as it goes, any music lover will object that music is more than architecture. A sensitive listener will be emotionally moved by good music, and we suspect that that is the primary reason for listening. While many composers and music experts claim that the structure is what listeners are responding to, few are specific about *how* the structure is related to emotional response. Attempts have been made to synthesize a conventional structure for a piece of music mechanically [Hiller & Isaacson 79], [Rader 74] but without any apparent reason other than to simulate music of a period or style. Form is analyzed with no mention of the purpose for choosing that particular form [Green 65].

Meyer presents a theory of meaning that does relate the structure of music to emotional response in the listener. According to him, the meaning of *meaning* in music is in the relationships between a stimulus (some sounds), that to which the sounds relate, and the conscious observer [Meyer 56, p. 34]. Music has meaning when it awakens expectations (meaning in relation to other music or within the music) in the listener.

2.3.4 Interestingness

The word "interestingness" is not found in our dictionary. It might be defined as the ability to arouse curiosity in an observer about an object or idea, or as a lack of "boringness." Meyer cites the law of Prägnanz which states that psychological organization will always be as "good" as the prevailing conditions allow. Thus a person will naturally try to find organization in everything he perceives. Reiteration gives a sense of unity, familiarity, expectation; a feeling of order; that the music is under control. A "wrong" or unexpected note, for example, engenders a strong yearning or expectation to "justify" it: resolve it somehow or integrate it into the music. Thus breaking a rule creates interest and emotion, but it must be justified or the listener will sense that the music is chaotic and not worth listening to. Once the mystery (tension) is resolved, the listener takes further instances of it for granted, and his interest moves on to the next mystery. Music, or rather the composer, "plays" on and with this principle. Good music holds the interest longer because its organization is not simply more complicated, but deeper, with many levels of complexity to be discovered and understood. Interesting compositions pose new and many-layered puzzles for the listener to solve. Boring ones pose no new problems to the listener who, having already solved $2 + 2$, has little interest in solving $2 + 3$. "Good" music should have an organization which, in hindsight "makes sense" in that later patterns result somehow from earlier ones, but in foresight cannot be reliably predicted ahead of time. The more patterns there are at more hierarchical levels, the more interesting the piece.

2.3.5 Aesthetics

Aesthetics or *beauty* is the trait in music that makes it not only "beautiful" but musical in the sense that it is recognizable as music. We believe that the deep purpose of beauty is to provide incentive to closer examination. This effect of aesthetics also occurs in science, where a kind of elegance or simplicity in a theory is considered "beautiful." In music, beauty is found traditionally in regularity (organization) and consonance, as well as some amount of chaos and dissonance to whet the desire for a return "home." "Consonance" in the previous sentence should be taken in the broadest possible sense - not only melodic and harmonic consonance, but rhythmic (regular meter and unsyncopated rhythms, for example), textural ("harmonious" groupings of instruments) and stylistic as well.

2.3.6 Tension and Release

The interplay of tension and relaxation is basic to all music and art, and is also related to mood, emotion and creativity. Basically, tension occurs when something unexpected happens (or is expected to happen but does not), or when it is not known what to expect. The amount of tension is directly related to the degree of unexpectedness. Relaxation (or release or satisfaction) occurs when the expected happens, even if it be undesirable (although relaxation is greater when the outcome is desirable). The amount of the release is directly related to the amount of expectation confirmed. Somewhere between the two lies boredom, when there is an extensive lack of tension, or (less often within the Classical style) there is an extended lack of satisfaction.

2.4 The Contributions of Deryck Cooke

Cooke's ideas deal with moods and how they are instilled in music. While his ideas are useful to this thesis, they do not form a "theory" in the sense that they seem to have no general or underlying principles. It is more a collection of observations than an attempt to derive universal rules. Cooke's ideas extend beyond what is needed here, so only that part used in this thesis will be discussed.

2.4.1 Motif

In this paper we will use *motif* to mean a list of about two to five "scale degrees" (pitches with reference to tonic but with no reference to particular octave or key) without any time relationship except that of order. A phrase (for the purposes of this thesis) is composed of *motive* — one to four motifs given specific rhythm and pitch — that are adaped in various ways by the moods' musical properties and Meyer's grammar according to the requirements of the input moods. Cooke examined many musical examples and distilled from them a group of "universal" motifs and their mood-implications. In the descriptions of these motifs in Table 2.1, numbers refer to diatonic scale degree (e.g., 1 is the tonic) and parenthesized numbers are optional or less important. The wording of the descriptions is mostly Cooke's.

The motifs are dominated by members of the tonic triad because the main key and mode-defining set of pitches should be stated early in the music to establish those attributes. The sixth scale degree plays a prominent role in many motifs, probably due to the fact that its quality is related to the mode. From this observation and Schenker's theory of hierarchical levels came the idea of the "skeletal

Table 2.1: Cooke's Motifs

1 (2) 3 (4) 5 - outgoing active assertive joy; exuberance, triumph, aspiration

5 1 (2) 3 ascending - outgoing emotion of joy, more pure and simple than above

(5) 6 5 harmonized (I)-IV-I – simple assertion of joy

- **alternation** – joyous vibration
- **6 as anacrusis** – with faint touch of longing
- **with 6 emphasized** – joyful serenity (if slow) with slight longing or pleading
- **6 as appogiatura** – burst of pleasurable longing

1 (2) (3) (4) 5 6 5 – innocence and purity of angels and children or birds

8 7 6 5 descending – incoming joy; acceptance or welcoming of comfort or fulfillment (consolation); feeling of continuing or never-ending joy

Major-mode Motifs

1 (2) 3 (4) 5 – outgoing feeling of pain, assertion of sorrow, protest against misfortune

5 1 (2) 3 ascending – pure tragedy, outgoing; firm and decisive — courage, heroism in the face of tragedy

5 3 (2) 1 arched – passionate outburst of painful emotion falling to acceptance; restless sorrow

1 (2) 3 (2) 1 arched – brooding, trapped fear; obsession with gloom; inescapable doom – context of immobility

(5) 6 5 as appogiatura to the dominant – burst of anguish

1 (2) (3) (4) 5 6 5 – powerful assertion of fundamental unhappiness

8 7 6 5 descending (natural minor) – incoming pain; acceptance of or yielding to grief; passive suffering; despair connected with death — feeling of open-endedness, never-ending despair

Minor-mode Motifs

motive" as a basis on which to build a phrase. There is also a consistent relationship between the contour of a motif and its mood description. For example, "incoming" motifs descend, and "sad" ones are in the minor mode. From these observations and Hevner's study on mood [Hevner 35] we suspect there may be a theory of motifs behind Cooke's observations.

2.4.2 The Intervals

Most music theorists (e.g., [Piston 62], [Schoenberg 67], [Hindemith 42]) agree that each member of a tonal (diatonic) scale has a particular "meaning," a tonal tension relative to the tonic and tendencies toward other scale-members. Cooke states them:

Tonic – neutral, finality; static

minor 2nd (m2) – finality; moves down to the tonic

Major 2nd (M2) – finality; moves down to the tonic (or up to the 3rd)

minor 3rd (m3) – fairly static (down to the tonic)

Major 3rd (M3) – fairly static (down to the tonic or up to the 5th)

Perfect 4th (P4) – moves down to the 3rd (especially a Major 3rd) but may move up to the 5th

TriTone (TT) – modulating to dominant key or up to the 5th

P5 (P5) – neutral, in flux, intermediate; moves up or down to the tonic (or 3rd)

m6 (m6) – in flux; moves down to the 5th

M6 (M6) – in flux; moves down to the 5th

m7 (m7) – no good resolution (down to the 6th or up to the octave)

M7 (M7) – finality; moves up to the octave

2.4.3 Vitalizing Agents

Vitalizing agents are used to embellish the basic motifs, to lend individuality and more precise meaning to them, and to transform them into playable music. We use the ones in Table 2.2 as the basis for the mood attributes.

Table 2.2: Cooke's Vitalizing Agents

Pitch -

- **contour** - upwards-moving implies more assertion, tension (compare with Cooke's ascending motif descriptions)
- **tessitura** -
  - **high** - implies lightness, transcendence
  - **low** - implies darkness, evil, heaviness

Repeated pitch -

- **slow** - repetitions represent monotonous deadness
- **fast** - repetitions represent continuous excitement without action; terrific drive

Rhythm -

- **rhythmic emphasis** - (a strong beat or a long duration) emphasizes the pitch
  - **duple rhythm** - feels rigid, controlled
  - **triple rhythm** - feels more relaxed, abandoned
- **evenness** - exudes smoothness, regularity
- **dotted rhythms** -
  - **fast ones** - display enormous tension and energy
  - **slow ones** -
    - **minor mode** - display weariness, solemness or tragedy
    - **major mode** - display lilting charm, pomp or courage

Tempo -

- **faster** - gives more animation to the music
- **slower** - makes the music serene or resigned

2.5 Meyer's Theories

2.5.1 Emotion

Emotion, according to Meyer, is generated by the basic emotional sequence:

1. Establish expectations, then
2. Inhibit or frustrate the expectations; finally
3. Fulfill or resolve the expectations.

The next few sections below will expand on these steps. Then some of Meyer's other ideas will be discussed. The ideas, while independent of the emotional "grammar" above, nevertheless support and further explain parts of it. The principle of saturation relates repetition to expectation. Reversal describes the effect of change. Completeness and closure show how fulfillment can be achieved.

2.5.1.1 Establishing an Expectation

"Expectation" of any kind must be based on prior experience or knowledge. There are two types of previous knowledge. Independent of any particular piece of music (but possibly dependent on period style) there exist some *general expectations.* They are learned from the listener's collection of previous musical experiences. *Specific expectation* is concerned with some particular musical *process* - a specific passage of music that has recognizable organization. ("Expectation" will often be used without the "specific" or "general" prefix since it is often not clear which word is appropriate.) A process, to be recognized as such, must be first established (i.e., stated) and then confirmed. The aspects of the process, such as its rhythm or contour, that are confirmed are assigned to a new sequence of notes — a "new" process developed from the original one by an appropriate method of elaboration. The confirmation may occur after a break from the pattern. A process is broken by changing or eliminating at least some traits of the process or introducing a new one. Clearly, the apparent break from or confirmation of a process can be relative. The strength of delineation of the break depends on the number of traits broken and the "strength" of each trait (this includes how long it has been in effect, how "obvious" it is, and how long it is expected to continue).

Tonality, meter, tempo and other "global" attributes of a piece of music tend to be established near the beginning of the melody. These attributes establish the mood as a set of global expectations about the music. The mood continues, remaining in the background, unless the global attributes are changed.

In addition to the initial establishment of the mood, a specific *pattern* upon which the subsequent content of the music is built is also usually established. This pattern is a grouping of notes that can be taken by the listener as a single unit.

Once it is perceived, the listener can become involved with the music (emotionally as well as intellectually) by following the *local* progress of the pattern and establishing some expectations about its destiny. Meyer's "Principle of Successive Comparison" states that the pattern will establish a local norm, a base for expectation (of precise repetition). Subsequent deviations from that pattern occurring in varied repetitions cause inhibition of expectation toward precise repetition. Tension arises out of these deviations from strict repetition. This stems from an innate desire in the listener for "order" or predictability, and frustration at any perceived disorder. The affect caused, according to Meyer, is very short-lived.

Confirmation is accomplished in EMOTER by continued establishment of the norms through reemphasis — especially of important notes — or slight modification. Although the change to the original motive may be slight, this step is necessary to the establishment of a "comfortable" phrase, especially in Classical music. In an emotionally-charged phrase, the grammar will direct that this step be left out to establish tension through uncertainty or ambiguity.

The specificity of an expectation produced by a passage depends on several factors: its completeness and well-formedness, whether a process is evident and evidence of a goal toward which the passage seems to be progressing. In a well-formed process, expectation is strong and specific for the process to continue to a specific conclusion. If a passage is well-formed but no process is yet evident, expectation is weaker. It is not certain if the passage is complete and ready to be confirmed with re-iteration or there is more passage to come before being made into a process. Expectation is unspecific except that something well-formed is to follow. With an ambiguous process expectation is strong to continue the process, but weak as to its goal, so expectation is also strong to change to a well-formed process. Saturation (explained below) controls the relative strengths of the conflicting expectations. Finally, with an ambiguous passage where no process is apparent, expectation is strong for a process or a well-formed passage (or both), but specificity for either is low, so almost anything new or different will lessen tension or resolve the passage.

The importance and meaning of a tone (or group of tones) is directly related to the strength and specificity of expectation generated by its occurrence. Below are some examples of expectation in music.

- An established process will tend to continue in like fashion.
- If such an established process is interrupted or changed, expect to return to it.
- In a disorganized or chaotic passage, expect organization to follow.
- If tension or dissonance increases, expect it to decrease (following the local climax).

- If a possible goal is perceived, expect to reach it (and do so "on time").
- Expect an active tone to resolve very soon.
- Expectations also arise from the position in time of a note group — e.g., no matter what the first few notes are, more will be expected.
- Groups tend to be of length 2^n, where n is a small integer.

2.5.1.2 Inhibiting Tendencies to Respond

Meyer's theory states that emotion is generated when a "tendency to respond is inhibited." That is, a listener's emotions are stimulated when his expectations are frustrated or go unfulfilled. The strength of frustration — and therefore the amount of emotion — is directly related to the length of time that the frustration was maintained and the "logic" of the expected resolution. Short range expectations are based on apprehending the method of *continuation* of a passage, while long range expectations are based on apprehending the method of *completion*. It is our opinion that emotion can also be generated *when that expectation is finally fulfilled*, or when it is perceived that the goal (fulfillment) is in "sight" — the climax or process reversal.

A change causes frustration; only exact repetition is completely without frustration. The frustration rules know several ways of altering or replacing a passage that will cause greater interest, emotion and frustration than the small changes possibly used in confirmation. Modification of a previous passage is the weakest; a change to something new is the strongest. Interestingly, the end of a phrase or melody also causes frustration. In the former case the amount of frustration can be controlled by the type (finality) of cadence, and by the amount of rhythmic and melodic continuation into the next phrase. In the latter case the end of a melody must be made as final and fulfilling as possible to overcome the frustrating aspect of sudden silence!

According to Meyer, a listener's emotions are also stimulated by an ambiguous passage. Such a passage will tend to produce two different expectations or no particular expectation at all. The listener will have little or no idea what is coming (or when) and his expectations will thus be inhibited. *Shape* is the relative amount of uniformity and diversity in a passage. Well-defined shapes (patterns or groups) require — and are created by — changes in pitch direction, and differences in duration and emphasis. Weak shape is caused by exaggeration of either differences or similiarities, but not both. Ambiguity is produced by a weak shape. Strong shape has a balance of similarities and differences.

2.5.1.3 Fulfillment

A listener's emotional response to inhibited expectation is based on the assumption that the music will eventually resolve frustrations to his satisfaction. The methods of fulfillment will be discussed in the section below on completeness and closure.

Inhibited expectations are resolved simply by continuing or completing the frustrated passage. Therefore, the method of expected continuation or completion must be known so that it can be eventually implemented. A more sophisticated system would also know *when* the resolution was expected to occur. A delayed resolution (if it is evident that the resolution is soon to be forthcoming) is more effective than one that occurs "on time."

2.5.2 Saturation

Another important principle Meyer recognizes is called *saturation*. It is based on the premise that too much repetition creates a desire for a change. Once the listener feels that he can easily predict the future course of the music based on its past, he has solved its mystery, and it no longer holds his interest. In addition, a deviation in the music, if excessively repeated, will weaken or cancel the effect of the deviation by transforming it into a norm (see tension and relaxation below). Therefore, as soon as a process is recognized (established and confirmed), a change begins to be expected, the expectation growing with each repetition. Expectation to change builds gradually, generally after three iterations (the first two being devoted to establishment and confirmation). Saturation may seem to be a contradictory idea in that there are expectations both to continue a repeating process and to change it (this is another reason why expectation to change is gradual with saturation). It must be remembered, however, that the listener becomes impatient when the music has too much sameness or too much chaos. This suggests that the number of iterations sufficient to cause saturation is inversely related to the well-formedness of the pattern members and to the simplicity of the method of iteration (repetition is simplest). The onset of saturation can be delayed by modifying the iterations.

2.5.3 Process Reversal and Climax

Reversal is the point at which a continuous process is broken and another mode of continuation takes its place. It is the climax and "turning point" of the passage, the point at which doubt and anticipation are replaced by more specific anticipation. The strength of a reversal is related to the amount of change, the "difference" across the reversal boundary (and the "width" of the boundary — abrupt or gradual).

2.5.4 Completeness and Closure

The mind strives toward completeness of particular shapes, regularity and simplicity of organization, and toward completeness based on style (e.g., resolve to the tonic, caesura) and context. Completeness is the fulfillment of an expectation, the finishing of something left unfinished. It is most effective after there has been a deviation from or frustration of an expected occurrence (*return*). While contiguous repetition increases tension (expectation of a change), return (after a digression) decreases tension. A return is more effective when finishing a process that was left incomplete than when repeating one that originally felt complete. Completeness and incompleteness of a note or note group can apply to different levels with a different value for each (e.g., a note can be a local tonic, but a global dominant).

Some examples of completeness include:

- A return to the original (beginning) harmony
- Finish ("fill") a pattern previously made incomplete (gaps or modifications)
- A (re)emphasis of norms, established patterns, etc.
- A varied recurrence (of a well-shaped term) that has strong "completion tendencies"

Some "destinations" toward completeness are:

- the Tonic (possibly local)
- a Strong beat
- a Return (to beginning)
- a Continuation (a very short-term destination)
- the Last note
- Completion of an incomplete repetition
- Loss of tension

Completion can be delayed (or made to seem incomplete) by:

- Applying an embellishment
- Repetition of an incomplete passage
- Frustrating expectations
- Failing to completely restate a group previously established as a unity (one or more gaps anywhere in the group)

Some examples of *in*completeness include:

- Uniformity of pitch, rhythm and meter; regular repetition
- *Part* of a pattern which (because of its history or its relation to a stylistic norm) is perceived as an identity - either a gap or unfinished (including delay in finishing)
- No meaning, no expectations, no process (or only a weak one)
- Rhythmic if first (or other locally strong) beat of a measure is omitted or made a rest — this explains the effect of syncopation

Closure occurs when a goal is reached. The length and strength of the goal itself after it has been reached is expected to be directly related to the "finality" of the passage containing that goal. Putting it another way, greater tension demands greater resolution.

Some examples of goals include:

- The melodic tonic
- An authentic cadence (dominant 7th chord to tonic)
- A point of relative repose (e.g., resolution of tension)
- Melodic descent
- Slowing down the
  - rhythmic activity
  - rate of musical process
  - sense of expectation
  - progress or development
- The end of a phrase, fragment or melody - a caesura near such a place creates the effect of a goal almost by default, whether or not the above is satisfied.

2.6 Conventional Music Ideas

In this section we will briefly discuss some "classical" music theory and philosophy of the type that has been commonly taught to music students. It is included to provide a basis for comparison with the theories used in this thesis. As stated above, classical theory considers musical meaning to be related only to music; almost no attempt is made to relate musical events to human response. Since the theory does seem valid and useful, however, we will try to justify some of it in terms of Meyer's theories. Philosophy considers musical meaning as a struggle between two opposing forces *within the music*. Again, little attempt is made to relate to the listener as Meyer does.

2.6.1 Ideas from Music Theory and Scholars

2.6.1.1 "Low-level" aesthetic rules (typical "music theory" rules)

Below are a sampling of good melodic progression rules quoted from [Palmer 72, pp. 48-49]. These are rules of thumb, rather than a strict body of theory, taken from observation of Classical music. Rules such as these have been used by [Hiller & Isaacson 79] and others in their music-writing programs. We have implemented a few of them to enhance the aesthetics of EMOTER's phrases.

- As a general rule avoid the leaps of a major seventh, all leaps beyond an octave and all augmented intervals.

- If a melody moves by a diminished interval let it return to a note within that interval.

- When the melody leaps a sixth or an octave, let the note preceding and following the leap be within it.

- As a general rule let the leading note rise to the octave, unless it is in the middle of a descending scale passage.

- A succession of large leaps is usually ugly.

2.6.1.2 Perceived Divisions and Grouping

Well-formedness or well-definedness or unity of a note-group (or note) is described by grouping rules in [Lerdahl & Jackendoff 83]. The lowest level of well-formedness of sound is a single note (actually, since grouping depends on contrast to delineate the groupingds, the space around the single note is a necessary part of the "lowest level of well-formedness"); therefore, a repeated note (rhythm and/or pitch and/or melodic interval) forms the simplest pattern. The delineation of (partition between) two groups is directly related to the contrast (of or) across the boundary (and the similarity of the groups). A musical passage can also be delineated from its neighboring passages by reiteration or by phrasing.

2.6.1.3 Schenker's Theories

Heinrich Schenker greatly influenced the music theorists of this century (including Meyer) with a new way to analyze music [Schenker 56]. Schenker believed that all great tonal music has a hierarchical structure with notes from the tonic triad as a "melody" at the highest level and the actual music at the lowest. A particular level is the result of tonal simplifications of the level below it and elaborations made to the level above.

Levels of Perception Schenker recognized three principal levels of musical perception: foreground, middleground and background. Each level abstracted away spurious details from the preceding one, leaving a more skeletal structure to study. This idea readily expands to allow many levels of perception. Another, thematically oriented system of levels is one used in this thesis. All levels have a rhythm (and pitch or pitch-group) — at the lowest level it is based on notes, at higher levels on larger and larger groupings of notes or patterns — motives, sections, phrases, etc. A note group may be "meaningful" on one level and not on another (higher) level. Therefore, different focus is necessary on different levels and many details can be ignored at higher levels.

Embellishment and Prolongation At each level music is expanded and explained by the processes of embellishment and prolongation. Simply stated, prolongation is a process whereby a tonality or chord is extended in time. A note is embellished by adding one or more auxiliary notes around it. Both techniques may be used to emphasize a note or chord, or to add variety to a passage. Either technique can and does operate on the other. Schenker's theory may be a good one for analysis, but it is not very helpful for synthesizing, since it never gives reasons *why* a certain embellishment or prolongation was chosen at a particular place in the music. The great majority of music theory is devoted to analysis, with the common belief that the processes that go on within the composer's mind are unknowable ([Jones 63], [Copland 57], [Toch 77], [Ericson 55]).

2.6.2 Tension-relaxation

Music can be thought of as a conflict between order and chaos where order wins out - at least in Classical music. The conflict is most interesting when the forces of chaos seem potent, but not overwhelming. This is true as far as it goes, but this says nothing about how the two sides are related. Conventional theory may say something like out of order grows chaos, and out of the chaos, order must crystalize, [Copland 57], but this is of little help when trying to explain music to a computer. Nevertheless, there is a constant interplay of tensions and relaxations. These forces can be compared with Meyer's notions of frustration and fulfillment of expectation. But Meyer goes beyond the dichotomy by explaining the cause of tension-frustration and relaxation-fulfillment:

- Relaxation - do what is expected:
  - continue a pattern
  - resolve a passage melodically or harmonically
  - restate or resume a pattern that was interrupted or changed or
  - end a phrase or section with a cadence or caesura (pause) after a decrease of activity and process
- Tension - frustrate expectations:
  - delay the expected (prepared tensions that are based on expectation created earlier in the music)
  - do the unexpected (e.g., high pitches, dissonances, fast jagged lines - these are unprepared tensions which frustrate general expectations) or
  - create ambiguity

After music attributes are discussed below we shall return to tension and release to see how they are related to the attributes.

2.7 The Meaning of Mood

2.7.1 Music Attributes

To be able to define moods (and manipulate music "intelligently"), the moods must be related to the music in some well-defined way. A musical passage has certain properties that we represent as a set of *attributes* chosen to be as complete and mutually independent as possible. Since only the melodic aspect of music is being considered, the attributes do not need to include such things as timbre, texture or voice-leading. Still there are many properties of melody, and unfortunately it is difficult to choose a fully-independent and complete set. The choices were carefully made so as to include what is commonly thought to be most important in writing a melody. They were based on how precisely a definition could be written, how much "meaning" could be attached, and how few attributes would do the job. They were also chosen on the basis of mutual independence of effect, so that they might be treated as independently as possible. It turns out that there are interactions among some of the attributes, but the original choices have for the most part been retained because they seem to be able to represent all the important properties of melody.

Below is a listing of the more important attributes with brief definitions. Chapter 5 covers them in more detail.

Tempo - number of beats per minute

Meter -

- **interbeat** - number of beats per measure
- **intrabeat** - beat division

Rhythm -

- **activity** - average number of attacks per beat
- **contour** - increasing or decreasing activity
- **unevenness** - variance of activity
- **syncopation** - a syncopated note lasts through a stronger metric division than the one on which it started
  - **interbeat** - fraction of beats that are syncopated
  - **intrabeat** - fraction of notes that are non-beat syncopated notes

Dissonance - strength and fraction of non-chord tones (strength is a product of metric location and duration)

weak - ratio of the number of unstressed dissonant tones to the total number of unstressed tones. These include passing-tones, grace-notes and neighboring-tones.

strong - ratio of the number of stressed dissonant tones to the total number of stressed tones. The only strong dissonance used in EMOTER at this time is the appogiatura.

Harmonic motion -

tempo - number of chord changes in the level

contour - difference between harmonic tempo at end and harmonic tempo at beginning

angularity - number of common tones between chords

naturalness - adherence to harmonic progression rules

Harmonic activeness - tone's function in the local harmony

Pitch -

important scale intervals (relative to tonic)

disjunctness - internote interval

scale activeness - fraction of active-note duration in the level

contour - location of the peak of the phrase

waviness - ratio of direction changes to number of notes

span - variance of pitches

Mode - major or minor

Melodic finality - strength of last pitch in the current (lowest) level

Phrasing -

length in measures

evenness - variance of lengths

beginning - part of measure

finality - metric strength of last note

Motive use - whether motive is reused, and how varied

Pattern -

strictness - exactness of repetition (rhythm carries more weight than pitch)

length of pattern-member relative to level

average number of iterations in the level

Motifs - Cooke's motifs and their meanings were used

2.7.2 Attribute Meanings

The most important property of the music attributes is their ability to imbue the music with personality. A particular mood has particular characteristics that can be captured as values of a subset of the attributes. The values for some mood could be listed here, but it may be more instructive instead to associate the attributes with tension and release.

2.7.2.1 Release Attributes

These are (our idea of) the "norm"; they describe a more or less average, normal melody. Of course, any motives or other patterns that become established through reiteration also become a part of the norm.

- Tempo – moderate (not slow)
- Rhythm – natural, fairly even, moderately active, unsyncopated; on strong beats
- Harmony – strong, simple, natural, tonal, consonant, slow "tempo"
- Range – moderately small
- Phrasing – even, natural, ending on last beat, ending on downbeat
- Pattern – long, moderate
- Melody – conjunct (repeated pitches); mild, falling contour
- Activity – low
- Pitches – tonic or tonic triad ending on tonic pitch

2.7.2.2 Tension Attributes

The tension attributes hardly need to be mentioned since they are the converse of the release attributes above. Any tension maintained for long enough, however, will be taken for granted and will become part of the norm. This explains why the style of Classical music has evolved from the elegant balance and relative simplicity of Haydn's time, to the darker complexity that is twentieth-century music.

2.8 Embellishment and Elaboration

2.8.1 Embellishment

Deciding which and how many embellishments to implement in a motive is preceded by the evaluation of all possible embellishments for that motive. The decision of whether or not to embellish the motive at all is based on how close to the desired attribute values the actual attribute values of the motive are. If, for example, the motive has an activity value at the maximum end of the desired activity value, no more embellishment may be needed. An embellishment E of a note N is considered a possible candidate if it meets certain requirements:

- There must be enough "room" for the embellishment. If N's duration is too short, E cannot be placed after it. Similarly, if the note preceding N is too short, no embellishment can be placed before N.

- Regardless of its duration, N must be an "important" note to allow it to be embellished. It must either be a chord-tone or an important scale degree. N should be metrically important as well. A grace-note cannot be embellished because whether or not it is consonant, it occurs on a very weak metric location.

2.8.1.1 Elaboration

The method of elaboration of a passage P (a motive or a phrase) is determined by P's emotional function (frustrate, fulfill, etc.), certain attribute values and by what has gone on before. Table 2.3 shows the permitted elaborations for each emotional function. The table was made using knowledge about the "meaning" or function of each kind of elaboration.

2.9 Summary and a Final Word

The body of theory cited above will, it is hoped, be sufficient to describe two qualities heretofore missing in computer-composed music: a sense of general mood that is specifiable, and a sense of purpose and goal. This is necessary if *musical* composition by machine is ever to be realized.

Before ending this chapter on theory, let us comment on the idea of "theory" itself. Basing a program intended to *compose* music on theories about music *analysis* has at least one problem. To use a travel analogy, there can be many ways to reach a particular destination, but once having arrived we can comment only on the route that was actually taken. Theory may "explain" a particular piece of music, but we should be able to construct innumerable other pieces which fit that explanation equally well. There is no single "greatest" masterpiece, and every

Table 2.3: Permitted Elaborations

Elaboration Type	confirm	frustrate	fulfill	ambiguate	process
repeat	√		√		√
transpose	√		√	√	√
reorder pitches	√	√		√	
embellish	√				
re-embellish		√		√	
remove a note		√			
generate new motive		√		√	
change a pitch			√		
slant				√	
use partial motive					√
syncopate				√	√
shorten duration				√	√

"perfect" composition could be rewritten in many different ways with no loss of its perfection or original meaning. This fact is both encouraging and discouraging, for although there are a plethora of solutions to any compositional problem, music theory alone does not tell us how to compute a single one of them! We must do some amount of searching through the maze of possibilities to find a musical one. EMOTER finds many "phrases," some of which we hope will be musical.

The next chapters of this thesis explain the methods used to represent the moods, emotions and music, and to implement the theory contained in this chapter.

Chapter 3

Representations

3.1 Introduction

Representing the creation, modification and storage of music is very much like modeling a composer's knowledge and thought processes. Just as the composer must know how to put the music down on paper, so must EMOTER have internally a precise, formal description of the notes and rules of music if it is to be able to compose. And just as the composer must know — at least on a personal level — what music means in order to compose expressively, so must EMOTER also have an extensive representation for this kind of meaning. Furthermore, not only must EMOTER's knowledge of these matters be stated explicitly in every detail, but it must be in a form useful to a computer. EMOTER's rules of music must be able to act upon the music; its knowledge about the meaning of music must be able to choose *which* rules to impose upon the music and when to impose them. These are the motivations and challenges we face when designing a system of representations for automated music composition.

We meet the challenge with the tools and techniques of Artificial Intelligence. Facts about music are organized and grouped by function and by their relationships with other facts. They are formatted in a uniform way so that they can be used efficiently. Their structure reflects their purpose. The rules that operate on these facts (and on other rules) also reflect in their structures relationships with similar rules.

Conventional music notation is a surprisingly dense storage method, but it is pictorial and therefore not computer-accessible. Exactly what information must the computer "know" and how can it be "known?" Obviously, there must be some method for storing the music — the notes themselves. Since generation of the music depends in this case on a set of moods as input the machine must also know what they mean and what to do with them. Music has a formal structure not directly reflected in the notes themselves; therefore a representation is needed for the structure. The attributes of music — tempo, meter, etc. — must be made accessible and understandable to the computer so that it can knowledgeably mold them to fit the desired moods. The many rules that constrain and operate on music must be represented in such a way that they can be used. This chapter discusses the representations chosen to implement the thesis. They are described in a declarative or procedural manner depending on whether they more resemble data structures or operations.

3.2 Representing the Music Itself

3.2.1 NOTES

Many computer representations for music exist (e.g., DARMS, EUTERPE), but all the ones we examined are unsuitable for EMOTER because they lacked the right combination of flexibility, clarity of notation and easy adaptability to a list-oriented programming language. We designed a very flexible, list-oriented representation for conventional (non-electronic) music called NOTES (NOTe Encoding System). NOTES is ideal for use with Lisp or PROLOG since it uses the list as its organizational method. The notation is easily readable and alterable by humans, and it can be expanded with new symbols to accomodate the user's needs. It can represent and store complex passages using whatever groupings of notes or parts seem most natural. It is not intended to represent any information about the physical placement of the music on the page. A more complete description of NOTES is presented in Appendix A. We originally intended to use NOTES in EMOTER, but we found we also needed an internal representation that would enable easy mathematical manipulation of the notes by the program. Also, since only melodic material was to be generated, with no consideration of explicit harmony (vertical chords), orchestration or many other factors that NOTES could handle, we decided that the sophistication, complexity and flexibility of this representation was more than we needed.

3.2.2 SICCLLL

A representation specifically for the present research was devised. It is called SICCLLL (pronounced "cycle"), for Simplified Internal Computationally Convenient Low-Level Listing. Essentially a stripped-down (and souped-up) version of NOTES, SICCLLL cannot in its present form store simultaneous pitches (chords or multiple lines of music), nor can it represent the finer points of music such as accents or dynamic markings. It does, however, enable quick and efficient manipulation of the notes of music — transposition, evaluation and many useful little "tricks" that make our job much easier.

3.2.2.1 Time

A passage in SICCLLL is represented by an ordered list of *notes*. A note is a list of two elements: time and pitch. Time is itself a list of two elements: location in the phrase and duration of the note. The smallest unit of time is the *tic*; location and duration are both expressed in tics. A tic, by the way, has no predefined relation to a beat, except that it is a simple fraction of a beat (e.g., $\frac{1}{4}$, $\frac{1}{3}$ or 1 beat). A phrase is assumed to start at location zero, but pickup notes are given negative locations. With this arrangement the first tic location of a measure, when divided by the number of tics in a measure, always has a remainder of zero. Furthermore, the same principle holds with beats, half-beats, half-measures and so on. The metric "strength" (explained in Chapter 6) of a particular location is thus easy to calculate, assuming no changes of time-signature throughout the phrase (a reasonable assumption for Classical music).

3.2.2.2 Pitch

Pitch is expressed as a positive integer. There are two scales of "normalized" pitches: diatonic and chromatic.

The diatonic scale is equivalent to the white keys on the piano; the chromatic scale includes all twelve pitches of the octave. The diatonic normalized scale allows operations strictly within the diatonic scale, such as diatonic transposition or measurement of scale intervals. The numeric difference between any diatonic pitch and the one immediately below it is always one.

The chromatic normalized scale allows operations using chromatic pitches. The numeric difference between any chromatic pitch and the one immediately below it is always one. A simple mathematical formula allows easy conversion between the two scales, although conversion from chromatic to diatonic assumes a "white" (natural) chromatic pitch. Surprisingly, this simple conversion takes care of the fact that diatonic pitches are not evenly-spaced in the chromatic scale. Middle C, the normalized major-mode tonic, corresponds to the chromatic pitch number 46 and the diatonic pitch number 27. The A below middle C (the normalized tonic for the minor mode) has a chromatic pitch number of 43 and a diatonic pitch number 25.

$$\textit{chromatic-pitch} = \text{ROUND}\ (\textit{diatonic-pitch} \times 12/7)$$
$$\textit{diatonic-pitch} = \text{ROUND}\ (\textit{chromatic-pitch} \times 7/12)$$

If $4 <= ((\textit{chromatic-pitch} \times 7) \text{ MOD } 12) <= 8$ then *chromatic-pitch* is a sharped or flatted note (a "black key" on the piano).

Below is a fragment of music[1] in $\frac{3}{4}$ time, the key of A minor and in the treble clef. Underneath it we display the equivalent SICCLLL notation, first with chromatic

[1]Not written by EMOTER

pitches, then with diatonic pitches. For this example a tic has a duration of one eighth-note (the smallest duration needed). The "-2" in the first SICCLLL note indicates that the first measure of the music is incomplete. The example begins 2 tics (a quarter-note) before the downbeat of the first full measure (which has a SICCLLL location of "0"). Continuing with the first note, the "2" following the "-2" gives the duration of the first note — two tics, or one quarter-note. After this parenthesized pair of time-related values is the pitch (corresponding to the musical pitch E), represented in the upper list by the chromatic value "50" and in the lower list by the diatonic value "29". The other SICCLLL notes follow the same rules.

The "*" in the diatonic version is *not* part of SICCLLL's notation. In fact, this particular musical example would generate an error if EMOTER actually tried to represent it using SICCLLL's diatonic pitch notation (there is no problem using SICCLLL's *chromatic* notation, of course). We show this to demonstrate that diatonic pitch representation in SICCLLL is limited to diatonic pitches.

```
((-2 2) 50) ((0 4) 55) ((4 4) 58) ((8 1) 57) ((9 1) 55) ((10 1) 54) ((11 1) 57)

((-2 2) 29) ((0 4) 32) ((4 4) 34) ((8 1) 33) ((9 1) 32) ((10 1) *) ((11 1) 33)
```

Occasionally one of the three parts is used for another purpose. A negative pitch is a place-marker when the program needs to "rotate" a motive. A duration of zero indicates that the note is a dummy that delimits a list of notes in which an end note is not yet determined, or a note to be joined with an adjacent one of the same pitch.

3.3 Representing the Music's Moods

The *mood* of a melodic passage refers to a general state of mind or feeling that is associated with it. The mood is assumed to stay constant within the span of a phrase (although in a longer piece it may change considerably). For this thesis we have chosen a representative set of moods, using [Dutch 64], [Cooke 83], [Lundin 53]. They are listed below, partitioned as well as possible into "axes":

Happy – Sad
Passive – Active

Static, Weary – Lively
Tense (Rigid) – Calm (Relaxed)
Protesting, Anguished (Longing) – Accepting
Outgoing – Incoming
Final – Continuing
Dark – Light
Relaxed – Rigid
Mild (Calm) – Passionate (Tense)

Each mood is represented by a (static) data base of attributes (described below), called a *moodframe*. These attempt to capture the characteristic attributes of the mood. For example, the most important moodframe elements for the *passive* mood are in Table 3.1.

In the actual program the moodframes are organized not by mood, but by attribute name, and kept in a file with the code for computing that attribute. Thus the *cadence* moodframe values for all the moods are in the CADENCES file along with the rules that calculate the cadence for a given set of moods. The "code" version of the example in Table 3.1 is in Table 3.2. A semicolon preceding a line indicates that the line is a comment for the preceding code.

3.4 Representing the Computed Attributes

Webster's dictionary defines an *attribute* as "a characteristic quality of a thing." The attributes for the phrase to be composed are computed from the moodframes of the input mood specification. For the present research these attributes are confined to properties a phrase possesses that can be manipulated to affect its mood. The decision to have a representation (or set of representations) for some of these properties was an important one made early in the development of EMOTER. Attributes seemed the natural way to implement theories about mood in music, but they also turned out to be of great consequence when the program needed to make decisions concerning the implementation of emotional expression. Each attribute (except motif and interval, which will be explained separately) is represented by a list that includes the name of the attribute and its values:

```
((attribute <attribute name> <list of value(s)>))
```

For example, the coded attribute for activity might be

```
((attribute activity ((melodic 0) (harmonic 1))))
```

Table 3.1: Important Moodframe Elements for the Passive Mood

- Motif – 8 - 7 - 6 - 5 (in C major, the pitches would be C - B - A - G descending)
- Important intervals – minor 2nd, 3rds, perfect 5th
- Tempo – not fast
- Meter – duple, simple
- Rhythm – even, natural
- Harmony – slow "tempo", small chord changes
- Cadence – plagal
- Phrasing – even, beginning on beat
- Pattern – repetition
- Melodic Contour – arpeggiated, reuse, small contour, moderately narrow span
- Activity – low
- Finality – moderate
- Dissonance – moderately low

Table 3.2: Encoded Moodframe Version of Table 3.1

```
    ((motif ( 46 45 43 41 )
;             8  7  6  5 - four ``standardized'' pitches descending
;                            diatonically stepwise from the tonic
        major falling
;        mode contour
(passive calm accepting incoming continuing relaxed happy mild)))

    ((melodic_interval (m2 (11 8))
        (meaning (passive weary sad anguished final dark))))
    ((melodic_interval (m3 (1 10))
        (meaning (passive weary sad dark))))
    ((melodic_interval (M3 (2 11))
        (meaning (passive calm happy mild))))
    ((melodic_interval (P5 (5 2))
        (meaning (passive active mild))))
; four melodic intervals are associated with this mood

    ((moodframe passive tempo 60))

    ((moodframe passive timesig ( 1     0     1     0     0 )))
;                                2/4   3/4   4/4   6/8   9/8

    ((moodframe passive harmony ((naturalness 3) (chromaticness 0)
        (tonicness 0.67) (tempo 2) (angularity 2))))

    ((moodframe passive cadences (0.5        0          0     1     )))
;                                authentic  deceptive  half  plagal

    ((moodframe passive phrasing ((length any) (naturalness any)
        (begin on) (end off))))

    ((moodframe passive pattern (0.5      2         )))
;                   length   strictness

    ((moodframe passive contour ((angularity 1.5) (span 2.5) (shape -1)
        (wavyness any))))

    ((moodframe passive activity (0 1)))

    ((moodframe passive finality 0.6667))

    ((moodframe passive dissonance (0           any     )))
;                                  unstressed  stressed
```

Motifs are represented differently. More than one motif is generally chosen to represent the mood specification. The name of each motif is simply its list of standardized pitches. In this context, "standardized pitches" means that a motif is represented as pitches in the octave belonging to middle C (a chromatic pitch of 46 as explained previously). The format for a motif attribute is:

```
((attribute motifs <list of motifs>))
```

with each motif having the format

```
(<list of standardized pitches> <mode type> <contour>
          <list of associated moodframe names>)
```

Below is a sample computed motif attribute.

```
((attribute motifs
  ((((53 50 46) major falling
    (accepting passive incoming happy relaxed final calm)))
   (((46 45 43 41) major falling
    (passive calm accepting incoming continuing relaxed happy mild))))))
```

Intervals are also represented differently:

```
((attribute intervals <ordered, evaluated list of intervals>))
```

Each interval in the ordered, evaluated list is represented by two integers denoting the chromatic interval in the major and minor modes, and the relative importance of that interval. The major and minor mode values are chromatic pitches located in the lowest represented octave (pitch values of 0 - 11). An actual example might be:

```
((attribute intervals (((2 11) 10) ((7 4) 7) ((5 2) 7) ((0 9) 7)
  ((8 5) 6) ((1 10) 6) ((10 7) 6) ((4 1) 5) ((11 8) 4) ((9 6) 3)
  ((6 3) 1) ((3 0) 1))))
```

3.4.1 An Example

In Table 3.3 is a sample set of moods and the attributes generated by them. For a more detailed description of the attributes, see the appendix on attribute representation and Chapter 4.

Table 3.3: Attributes for the Moods (happy calm passive)

```
((attribute activity ((melodic 0) (harmonic 1))))
((attribute cadences ((authentic 2.5) (plagal 2) (half 1)
  (deceptive 0))))
((attribute contour ((angularity 1.4166) (span 3.3333) (shape -0.5)
  (wavyness 8))))
((attribute dissonance ((weak 0.25) (strong 0))))
((attribute emotion ((value 0.3333) (slope 0))))
((attribute finality (0.8333)))
((attribute harmony ((naturalness 3) (chromaticness 0.0833)
  (tonicness 0.5566) (tempo 2) (angularity 1.6555))))
((attribute intervals (((2 11) 10) ((7 4) 7) ((5 2) 7) ((0 9) 7)
  ((8 5) 6) ((1 10) 6) ((10 7) 6) ((4 1) 5) ((11 8) 4) ((9 6) 3)
  ((6 3) 1) ((3 0) 1))))
((attribute pattern ((length 0.7071) (strictness 2.5))))
((attribute motiveuse (sequence 1.25)))
((attribute motifs
  ((((53 50 46) major falling
    (accepting passive incoming happy relaxed final calm)))
   (((46 45 43 41) major falling
    (passive calm accepting incoming continuing relaxed happy mild))))))
((attribute phrasing ((length 8) (naturalness 3) (begin on) (end on))))
((attribute rhythm
  ((activity (average 1.0606) (variance 1.8333) (contour 0))
   (shape (syncopation (interbeat 0.3333) (intrabeat 0.3333))
          (dottedness (interbeat 0.8818) (intrabeat 0.75))))))
((attribute tempo (93.3333)))
((attribute timesig ((4 4) 3)))

((normtimesig 4 2 8))
((measure_hierarchy (8 4 2 1)))
((metric_hierarchy (2.8284 2.0 1.4142 1)))
```

3.5 Representing the Motive Embellishments

A *skeletal motive* is a motif[2] that has been assigned a rhythm in accordance with the principle of rhythmic emphasis outlined in Chapter 6. Thereafter any changes made to the motive are called *embellishments*. Each type of embellishment has a set of rules that implement it, just as there are rules that give the motif a rhythm. In evaluating, selecting and implementing embellishments, each one is represented by its name and possibly one or two arguments. Most embellishments are associated with a particular motive note and occur just before or after the note. None include explicit rhythmical information, but a couple give some guidelines. Listed in the appendix on development methods are descriptions of the embellishments. Chapter 4 describes the process of embellishment selection. The embellishments used in EMOTER are the passing-tone, neighboring-tone, appogiatura, escape-tone, arpeggiation and repeated-tone.

3.6 Representing the Elaborations on Motives

In the same way that embellishments operate on notes in motives, *elaborations* operate on motives in a phrase. The word *passage* will be used to mean a contiguous group of notes. Each elaboration is supported by a set of rules and has a similar type of representation to an embellishment. (The self-similar or fractal nature of music at different different levels is evident here as it is, for example, in metric divisions and formal structure. Recognition of this property as a type of pattern facilitates the expression of a representation.) Although some of the elaborations may seem to lack musical justification, all are commonly used, if not explicitly labeled as such. This representative (but by no means exhaustive) set of elaborations should be sufficient to "express" the manipulation of the listener's expectations according to [Meyer 56] to produce an emotional response. Listed in the appendix on development methods are descriptions of the elaborations. Chapter 4 describes the process of elaboration selection. The elaborations used in EMOTER include

- repeating a passage
- transposing a passage
- reordering some of the pitches
- removal of a note
- changing a pitch

[2]Recall from Chapter 2 that a *motif* is a list of about two to five "scale degrees" (pitches with reference to tonic but with no reference to octave or particular key) without any time relationship except that of order.

- embellishing or reembellishing a passage
- generate a new motive
- syncopating part of a passage

3.7 Representing the Phrase Structure

To enable the phrase to have an internal organization and goal, its structure — the arrangement,groupings and relationships among the notes — must be represented somehow for reference and modification of motives composed later in the phrase. EMOTER goes through several stages of representation, described below.

The first and simplest stage, called PhraseStructure1, represents each of the motifs to be used in the phrase as an ordering-label number-pair. The first motif is given the ordering-number one, the next *different* motif (that is, not just the same motif repeated) is given the ordering-number two and so on. The first occurrence of a particular motif receives the label-number one, the next occurrence of that motif is given two, and so on. A phrase composed of motifs a,b,a,c would be represented in PhraseStructure1 as ((1 1) (2 1) (1 2) (3 1)).

The second stage, PhraseStructure2, stores the Meyer emotion grammar that will be imposed on the phrase. Each motif will either be ESTABLISHed, CONFIRMed, FRUSTRATEd or FULFILLed; the appropriate label is associated with the PhraseStructure1 number-pair. One possible assignment to the example in the preceding paragraph might be ((ESTABLISH (1 1)) (ESTABLISH (2 1)) (CONFIRM (1 2)) (FRUSTRATE (3 1))).

The third stage of phrase structure is the determination and subsequent representation of the melodic contour (and consequently the location of the climax of the phrase). To each motif is added for this purpose a number representing the length it will finally have in the phrase and its final pitch-placement (average pitch and harmonic content). These numbers follow the motif as it goes through its transformations on the way to becoming a fully-developed motive in the final phrase.

The actual embellishments and elaboration performed on a motive are not stored as a representation at this time. If complete melodies were being composed, that would be necessary in order to preserve the entire internal organization for later reference and manipulation; for this thesis it is sufficient to store the embellishments in SICCLLL and let an elaboration operate on an entire embellished motive. Thus the organization of the phrase is represented in part directly as notes and in part indirectly as instructions.

3.8 Summary

Representing the important aspects of the music in EMOTER requires many specialized techniques, from purely numerical to purely symbolic. We have seen in this chapter how each representation is specifically tailored to the needs and structure of what it is to represent. The operations and rules relate the representations to one another in a large network thatpossibly provides a simplified model of the mental processes of some human composers.

Chapter 4

The Overall Plan

4.1 Introduction

EMOTER consists of several major modules:

INPUT MOODS
⇓
PROCESS MOODS
⇓
COMPOSE MOTIVES
⇓
COMPOSE PHRASE
⇓
OUTPUT PHRASE

In brief, the user inputs a set of moods to EMOTER that he wants the phrase to "express." The moods are processed to obtain attribute values with which to direct the program's composition mechanism. Then one or more motives, which form the thematic source of the phrase, are generated using the attribute values and a set of embellishment rules to help make decisions. The phrase is built from the (elaborated) motives according to Meyer's theory of emotion.

The modules are described in more detail below. To clarify the explanation, the operation of EMOTER will be traced in an example. We will use the following phrase (also shown in the introduction).

4.2 Input of the Moods

Input is very short and simple. First, EMOTER requests a title for the phrase. The user is then prompted for a list of moods (enclosed in parentheses for ease of processing). The program verifies their spelling and asserts them (later they will also be examined for "impossible" combinations). Upon request the program will supply a list of the valid moods. Below is a sample input session, showing the title of the phrase to be composed and a list of the desired moods for the phrase. Prompts from the program are not shown.

```
Angst Phrase
(outgoing protesting tense active)
```

4.3 Mood Processing

To be useful (and comprehensible) to EMOTER, the moods must be translated into terms EMOTER can understand, and with which it can manipulate the music. To accomplish this, EMOTER has stored a special definition of each mood called a *moodframe*. The moodframes for the input moods are collected, processed and combined to form a set of *attributes* that completely describe the input moods as a gestalt to the program's composing mechanisms. The computed attributes are everywhere visible to EMOTER's routines and, as we shall see, exert an influence over nearly all of them. Derivation of the attributes' values is covered in Chapter 5. Here we will only give the computed attributes in Table 4.1. The table should be carefully examined since it will be referenced in this and following chapters as the trace unfolds. In particular, notice that

- The contour angularity is near its highest value, indicating that the phrase will tend to have disproportionately many melodic leaps.
- The contour span is also high. The phrase will therefore span considerably more than an octave.
- The contour shape indicates that the peak or climax of the phrase will occur near the end of the phrase.
- (Rhythmically) weak dissonance is at its highest value, indicating the likelihood of such embellishments as escape tones and neighboring tones (and the *un*likelihood of consonant embellishments).
- Emotion is quite high on the average, and the positive slope indicates that the emotional level will increase.

- The most important scale degrees (from the `intervals` list) are the fifth and tonic. This information will come into play when choosing the skeletal rhythm for the motifs.
- Pattern length indicates that the length of the shortest motive will be about one-quarter of the total length of the phrase.
- Motiveuse calls for a contrasting set of motifs rather than, say a phrase using only one motif.
- The two motifs to be chosen first are (43 50 51 50) and (43 46 50). These are minor-mode motifs. In scale-degree terms, they are (1 5 6 5) and (1 3 5).
- Phrasing calls for the phrase to begin off the beat, which means that the phrase will begin before the downbeat of the first full measure.
- Rhythm activity indicates that there are to be two or three notes per beat on the average.
- The three attributes at the bottom of the list (separated by a blank line) are the only ones that do not have corresponding pre-existing moodframes. They are computed instead from other attributes and from EMOTER's knowledge about rhythm and meter.

We will see how these attributes come into play in a number of ways during all stages of the composition process.

4.4 Composition of the Motives

4.4.1 Introduction

Before beginning a detailed examination of the single most complicated and important part of EMOTER, we first prepare the reader with some terms and guideposts.

4.4.1.1 Evolution of a Motif

Before starting into this section, let us distinguish among four stages a group of notes must go through before becoming a phrase. Recall that a *motif* is a series of scale-degrees (in a particular mode but without actual assigned pitches) having no rhythmic attributes. A *transitional motif* is a sequence of notes which is on its way to becoming a motive, but is not yet one. A *skeletal motive* has not yet been embellished or otherwise developed. A *motive* has specific pitches, a metric location and a rhythm. It may or may not have been embellished. A *phrase-member* is a generic term meaning any one of these.

Table 4.1: The Computed Attribute Values

```
((activity ((melodic 1) (harmonic 3))))
((cadences  ((half 2) (deceptive 1) (plagal 0) (authentic 0))))
((contour ((angularity 3.7) (span 6.7) (shape 0.5) (wavyness 0.6))))
((dissonance ((weak 1) (strong 0.5))))
((emotion ((value 0.8) (slope 0.5))))
((finality (0.2)))
((harmony ((naturalness 1) (chromaticness 0.5) (tonicness 0.2)
           (tempo 2.5) (angularity 1.6555))))
((intervals (((5 2) 10) ((10 7) 10) ((4 1) 7) ((8 5) 6) ((9 6) 5)
             ((7 4) 5) ((6 3) 5) ((3 0) 5) ((2 11) 4) ((1 10) 4)
             ((0 9) 3) ((11 8) 2))))
((pattern ((length 0.251) (strictness 1))))
((motiveuse (contrast 3.5)))
((motifs (
*   (((43 50 51 50) minor arched
*        (tense lively active sad protesting dark passionate))
*     ((43 46 50) minor rising
*        (tense active sad protesting outgoing passionate)))
    (((38 43 46) minor rising (outgoing passionate sad))
     ((43 46 50) minor rising
        (tense active sad protesting outgoing passionate)))
    (((51 50) minor falling
        (tense lively active longing anguished passionate sad))
     ((43 46 50) minor rising
        (tense active sad protesting outgoing passionate)))
    (((43 46 43) minor arched
        (tense static incoming dark passionate sad))
     ((43 46 50) minor rising
        (tense active sad protesting outgoing passionate))))))
((phrasing ((length 3) (naturalness 2.9661) (begin off) (end off))))
((rhythm ((activity (average 2.4494) (variance 2) (contour -1))
          (shape (syncopation (interbeat 1) (intrabeat 1))
                 (dottedness (interbeat 0.1) (intrabeat 0.1))))))
((tempo (100)))
((timesig ((2 4) 1)))

((normtimesig 2 4 8))
((measure-hierarchy (8 4 2 1)))
((metric-hierarchy (1.4142 1 0.7071 0.4999)))
```

4.4.1.2 An Outline of the Motive-Composition Process

To compose the motives, the following major steps are performed:

1. A list of motifs is chosen and ordered.
2. The phrase structure is determined.
3. The lengths of the finished motives are calculated.
4. A contour is computed for the phrase.
5. The motifs are assigned skeletal rhythms.
6. The skeletal motives are embellished.

4.4.2 Initial Choice and Ordering of the Motifs

From the chosen set of motifs (the first set — preceded by asterisks — in Table 5.1) all "legal" motif-combinations are enumerated, then ordered. Hereafter these motif-combinations will be called *skeletal phrases.* The number of motif-instances is determined by the pattern length attribute and the constraint that each motif must appear in the phrase at least once. In the present example, the pattern (motive) length value is about one-quarter (0.251) of a phrase, so there should be about four (3.984) motives in the phrase. The legality of a skeletal phrase is determined by pattern length (number of *motives*) and *motiveuse*. The number of consecutive iterations of the first motif should be approximately equal to

$$motives - ((motiveuse \div 4) \times (motives - 1)).^1 \qquad (4.1)$$

If not, the entire skeletal phrase is discarded. If it passes this test, the rest of the motifs in the chosen skeletal phrase are examined similarly, using a new value of *motives* equal to the previous value minus the number of consecutive iterations of the first motif. For this example there should be about

$$3.984 - (3.5 \div 4 \times (3.984 - 1)) \cong 1.65$$

(i.e., two or one) consecutive iterations of the first motif.

Each legal skeletal phrase is evaluated using motive-use and the following method:

1. Let TEMP equal the product of the numbers of identical motifs in a row, times (four divided by the string-length), divided by the number of unique motifs.

[1] *Motiveuse* is divided by four because, as the reader may remember, its value ranges from zero to four. The idea is that the smaller its value (the more repetitiousness which is called for), the more consecutive repetitions of the same motif should be encouraged.

2. Find the numeric difference between TEMP and (four minus *motiveuse*).

3. The value of the skeletal phrase is one minus ((the absolute value of the above difference) divided by four)[2].

This evaluates the closeness of the consecutive iterations in the skeletal phrase to the desired motive-use and provides a ranked ordering among the candidates.

The skeletal phrases which have passed the above constraints are now examined in the light of Meyer's emotion grammar. Unless motive use expects high contrast in the phrase (as is the case in our example), a skeletal phrase with an initially repeating motif will be preferred over one that starts with two different motifs. The most-preferred skeletal phrase from the above constraining processing is chosen as the basis of the phrase. The one chosen for our example is (1-5-6-5, 1-3-5, 1-5-6-5) and is shown below.

(4.2)

Its first motif, 1-5-6-5, we label **1**. It has an arched contour and is in the minor mode. Its associated moods are tense, lively, active, sad, dark, protesting and passionate — three of which fit the input mood specification of outgoing, protesting, tense and active. The other motif chosen, 1-3-5, we label **2**. It has a rising contour and (since all motifs must be in the same mode) is also in the minor mode. Its associated moods are tense, active, sad, protesting, outgoing and passionate — and all four input moods are members of its list. Were it not for the requirement from the motive use attribute for contrasting motifs, **2** would have been the sole motif of this skeletal phrase.

4.4.3 Determination of the Phrase Structure

From the ordering of the motifs and Meyer's emotion rule (see Chapter 2) a *Meyer-phrase-structure* is chosen for the skeletal phrase. The choice is made using a small set of rules that have knowledge about the relationship of motif-order, emotion, and Meyer-phrase-structure built into them. A description of the process follows.

First we determine the basic structure of the phrase (PhraseStructure1, discussed in chapter 3). In our example it is ((**1** *1*) (**2** *1*) (**1** *2*)), meaning that a

[2]As before, the fours in this method come from the range of values of *motiveuse*.

motif arbitrarily numbered **1** (different motif numbered **2** (1-3-5) (occurring only *once*), followed by a *second* iteration of motif **1**.

From the above basic structure and the emotional contour, the most Meyer-like emotion-progression is decided (PhraseStructure2, also from Chapter 3). There is nothing particularly sophisticated about these rules; they simply match up the basic structure with the desired emotional slope attribute and return a Meyer-emotion sequence. This knowledge was imparted to the rules in a direct fashion by enumerating the possibilities and constructing the rules based on our interpretation of Meyer's theory. In our example the slope is positive (thus the emotional level is to increase in the phrase); this combined with the basic structure will result in an emotion-progession of the ESTABLISHMENT of motives **1** and **2**, followed by FRUSTRATION of motive **2**. The actual code produced is

```
((ESTABLISH (1 1)) ((ESTABLISH (2 1)) ((FRUSTRATE (1 1))).
```

Since there is no fulfillment in this phrase, it obviously could not be the last phrase of a melody[3], but it should be an appropriate first phrase of a highly emotional one.

4.4.4 Determination of Phrase Contour and Harmonic Progression

Now that EMOTER knows which motifs are to be used, the order they are to appear in and what general effect the Meyer-emotion grammar is to have on them, the dimensions of the phrase containing them (the motive lengths, melodic contour and harmonic context) can be determined. Using the computed attributes for melodic contour, harmonic tempo, normalized mode and general knowledge of the intended time-placement of each motive in the phrase, the vertical (pitch-area) and horizontal (time-area) placement for each phrase-member is calculated. The actual adjustments are not made until much later, however.

4.4.4.1 Determination of Lengths of the Motives

Possible lengths for the motives are evaluated first (before the pitch-locations), based on the phrase length, the constraint that each implementation of a particular motif must be the same length, and the constraint of motive length to a whole number of measures (or a simple fraction of a measure for a very short phrase)[4].

[3] As EMOTER is written presently it has no knowledge of entire melodies. When it is expanded to compose melodies this will be one kind of knowledge it will be given.

[4] The actual implementation of the length-constraining rules was as a knowledge-base, exhaustively listing every allowed combination of motive lengths.

Beyond that, motive lengths proportional to the number of pitches in each motif are preferred. With the sample phrase length previously chosen to be four measures and a **1-2-1** phrase structure, the only possible motive lengths are one measure for the **1** motives and two for the **2** motive[5].

4.4.4.2 Calculation of Contour and Harmony

Now that each motive's length has been computed, a melodic contour and chord progression for the phrase can be constructed. EMOTER here makes a simplifying assumption that each motive has just one harmony[6]. First, we will discuss how an average approximate pitch-location is calculated for each motive. Then we will show how its exact pitch-location is "fine-tuned" to conform to the constraints of harmonic progression and motive length.

Phrase Contour Calculation The pitch-locations of the motives are "roughed-in" by placing them along the contour of the phrase. The contour itself is calculated from the emotion and contour attributes and the chosen phrase length using geometry. Figure 4.1 represents a generic phrase. The height (called "span" in the contour attribute), phrase length and general shape are supplied by the attributes. The peak's metric location is calculated from the shape (relative time-location of the peak in the phrase) and phrase length. The (diatonic pitch) height is calculated from the span attribute and phrase length. The exact start and end pitches (and therefore the exact peak pitch) are not yet known.

Calculation of the Pitch Areas of the Phrase Members EMOTER now finds an appropriate transposition interval for each motive. (Note that the motives will not be actually transposed until they have been completely composed otherwise — that is, embellished and elaborated.) It does this by choosing the exact start and end pitch areas, and from these the peak pitch. The first and last motives are assigned the two end pitches as their transpositions; any remaining motives are placed between them along the contour defined by the three pitches in a manner to be described.

EMOTER calculates these three pitch "vertices" of the phrase by first determining how the phrase should begin and end emotionally. The two emotion attributes, level and slope, together describe a general emotional contour for the phrase. Here we are using the word "emotion" in its one-dimensional-axis sense,

[5]Because physical passage lengths are not part of any theory we are using, it was necessary to make some simplifying assumptions for this section. These assumptions do, however, have their origins in common practices of Classical music composition.

[6]As usual with the simplifying assumptions, this one was made because none of the theories we are implementing cover this problem and because in Classical music practice a motive instance generally is associated with only one harmony.

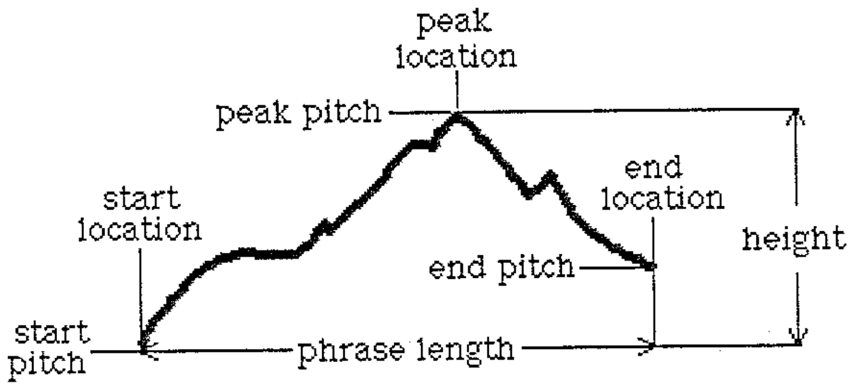


Figure 4.1: A Sample Melodic Contour

not to refer to Meyer's idea of emotion progression. Our example has a calculated emotional level of 0.8333 (out of a maximum emotion of 1) and a slope of 0.5, indicating that the emotional level is to increase during the phrase. For contour, the starting and ending emotional level is represented musically by a particular choice of chord for the first and last motives: tonic for low emotion and dominant for high emotion (for an entire melody these two would be increased to include all harmonies).

In our example, the first chord is chosen to be tonic and the last to be dominant. But should the dominant be a fifth above the tonic or a fourth below it? The contour-shape attribute, which in our example places the peak near the end of the phrase, dictates that the former be chosen. Now the height of the peak (above the lowest point – the first motive) can be determined from the contour span attribute, and the time- and pitch-location of the lowest motive. The peak here turns out to be about two octaves above the lowest pitch of the phrase (but no harmony is implied at the peak). The outline of the phrase contour has now been calculated (and the end motives' transpositions determined).

Harmonic Progression Calculations The positions of the middle phrase member(s) are derived from the phrase contour outline and EMOTER's knowledge about legal chord-progressions. The approximate pitch-location of each middle member is determined again from geometric considerations — if one considers it to be a point, it should fall near the "line" of the contour, at its average (central) time-location. The chord for the middle motive in this example was chosen to be a tonic like that of the first, but because of the steep contour the second is placed an octave above the first. The three "transitional motifs" are finally ready to be individually shaped into true motives and joined together. However, at this point *no changes to the transitional motifs have been made.* They are still in "root position" (tonic triads with no rhythm, centered around middle C or A). The information just computed (diatonic transpositions of 0, 7 and 4 diatonic steps (upward) for the first, second and third motives, respectively) is merely carried along with them for now.

4.4.5 Motive Composition – First Time

The transitional motifs are now assigned specific rhythms and metric placements, embellished and further developed into a phrase. An original motive is *composed* only once — thereafter it is developed, or *elaborated.* This "first-time" composition thus establishes the thematic material, and each reappearance of that material is expected to be related to it. The plan of composition of this material is outlined below. The operations performed on a phrase-member depend on its Meyer-emotion grammar function in the phrase:

ESTABLISHment consists in simply "composing" a motive out of a transitional motif (and then transposing it to its intended pitch-location);

CONFIRMation requires a pre-existing motive to be confirmed. That motive (used in "tonic" position to simplify working with it) can be repeated, slightly elaborated or re-ESTABLISHed;

FRUSTRATion of a motive is similar to confirmation, but with more radical changes made to the motive;

FULFILLment is also similar to confirmation, with only the most stable operations allowed on the motive.

Since two phrase-members are to be ESTABLISHed we need trace the progress of only motif **1** (1-5-6-5), which will later be FRUSTRATEd. Some interesting operations on motif **2**, however, will also be examined.

4.4.5.1 Establishment of the First Motive

As mentioned above, motif **1** is "composed" to produce motive **1**. The process of motive composition consists of assigning a skeletal or basic rhythm to the motif, then embellishing the skeletal motive. We give a detailed description of the skeletal rhythm composition process (including a trace of our example) in Chapter 6. For the sake of continuity in this chapter the resultant skeletal motive is shown below.

(4.3)

```
(((43 7) 0.7071) ((0 0) 0.5) ((50 10) 1.4142) ((0 0) 0.5)
((51 8) 0.7071) ((0 0) 0.5) ((51 10) 1.0) ((0 0) 0.5))
```

Final Touches to the Skeletal Motive Now all pitches have been rhythmically placed at the strongest possible locations. Each note has both a duration and an absolute location. The pitches are still "normalized" to the tonic harmony — the motive is still carrying the number representing a diatonic transposition interval with it.

At this point a "dummy" last note is added to the motive-group for convenience in executing the next step — embellishment of the skeletal motive. The note has the same pitch as the first note in the motive-group, but its duration is set to zero to mark it so that it can be easily found for removal later. As we shall see, every embellishment operation is carried out between two adjacent notes. The dummy note provides a temporary "partner" for the real last note so that it too can have an opportunity for embellishment.

The skeletal motives in our example are shown in Figure 4.2 (in SICCLLL notation). The vertical lines signify bar lines.

Notice the eighth-note gap between the two underlined notes, starting at tic number 6 with a duration of two tics. This will be filled in later when the finished phrase members are finally put together.

4.4.5.2 Embellishment of the Motive

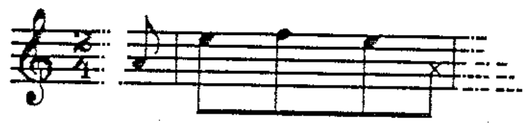
The final operation in composing an original motive is *embellishment*. The skeletal motive is embellished by appropriate insertion of auxiliary notes into the passage. Embellishment is done to

- emphasize certain notes (specifically, the most important ones) in the skeletal motive over others by "decorating" them to make them stand out;
- implement or express certain mood attributes (e.g., syncopation, dissonance, melodic angularity), thereby helping to express the desired moods;
- add interest, creativity and beauty to what would otherwise be a bare frame of a motive; and
- make the motive unique to the particular piece of music.

The process of choosing and implementing the embellishments is covered in Chapter 7. Here we merely display the end result of embellishment for our example. Lest the reader be led to believe from the above list of justifications that embellishment of a skeletal motive is *required*, he should note that EMOTER has decided *not* to embellish motive 1 at all: this skeletal motive remains as it was. Motive 2, however is liberally embellished as can be seen by comparison of its skeletal version in Example 4.5 with the embellished version in Example 4.6. The numbers in boxes are measure numbers.

(4.6)

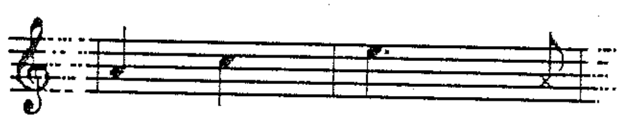
Skeletal Motive 1:



(4.4)

```
SICCLLL        (((-2 2) 43) ((0 2) 50) ((2 2) 51) ((4 2) 50) ((6 0) 43))
                            |                                 -----------
Locations        ``and''    | ``One''    ``and''    ``Two''    ''and''
Durations         eighth    | eighth     eighth     eighth       0
Scale Degrees          1    |       5          6          5   dummy note
```

Skeletal Motive 2:



(4.5)

```
SICCLLL          (((8 4) 43) ((12 4) 46)   ((16 6) 50)         (22 0) 43))
               |  ----------             |
Locations      |  ``One''    ``Two''     |  ``One''(``two'') ``and''
Durations      |  quarter    quarter     | dotted-quarter       0
Scale Degrees  |        1          3     |        5          dummy note
```

Figure 4.2: The Skeletal Motives

4.5 Phrase Composition

To finish composition of the phrase the following steps are performed:

1. The embellished motives are elaborated to fill out the rest of the phrase, guided by Meyer's emotion grammar.
2. The seams between the motives are smoothed.
3. Any necessary compensation for the minor mode is made.

4.5.1 Elaboration

Recall that a phrase is made up of one or more motive-instances. The relationships among them are determined by Meyer's emotion grammar. Implementation of Meyer's grammar is accomplished in EMOTER by motive elaboration. A first-time motive is always *established*, which is to say that no elaboration as such is performed upon it. Instead it is a composed and embellished motif as described above. When a motive reappears in a phrase it is then elaborated by being CONFIRMED, FRUSTRATED or FULFILLED, depending on the requirements of Meyer's grammar. Each of these grammar-operations has associated with it some particular lower-level elaborations such as transposition or syncopation. The particular elaboration chosen for a motive is based on such attributes as emotional level and contour, pattern strictness, motive use and the harmonic attributes (not to mention Meyer's grammar). Elaborations range from strict repetition (good for confirmation of a motive) to the composition of a completely new motive with no thematic relation to what has preceded it. For more complete information about the elaboration process and an example of that process, see Chapter 7. In our example the third phrase-member is to be a FRUSTRATEd elaboration of motive 1 (1-5-6-5). For reasons explained in Chapter 7 the elaboration chosen was the removal of an important note (the first "5" from 1-5-6-5, leaving 1—6-5) from motive **1**. To see the effect of the elaboration, compare the original motive (left) with the elaborated motive **1** in Example 4.7.

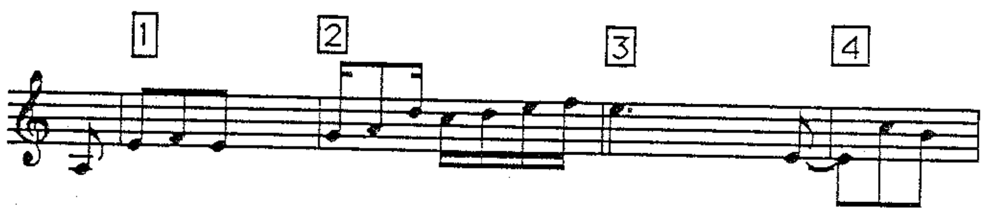
(4.7)

4.5.2 Interfacing the Phrase-members

Once all motives have been fully embellished or elaborated (and transposed to their proper pitch-areas), they must be connected together to form a seamless phrase. The rules EMOTER uses for connection are the following (in this order):

1. Motives separated by too great a leap (more than an octave) must be brought together by transposition of one toward the other by exactly an octave, to retain the original harmonic progression.
2. Any gaps between motives (places where there is no note at all) must be filled. Any embellishment may be used; the notes on either side of the gap are treated as a note-pair for the embellishment routines.
3. Finally, to smooth the transition between motives, passing-tones may be used if necessary (here the embellishment mechanism is also applied).

Here is our sample phrase before its motives have been interfaced. The blank lines in the SICCLLL notation separate the three phrase-members.



(4.8)

```
(((-2 2) 43)
 ((0 2) 50) ((2 2) 51) ((4 2) 50))

 ((8 1) 53) ((9 2) 55) ((11 1) 60)
 ((12 1) 58) ((13 1) 60) ((14 1) 62) ((15 1) 63)
 ((16 6) 62)

                ((22 4) 50)
 ((26 2) 58) ((28 2) 57))
```

The transition between the second and third phrase-member does not offer quite enough space for six passing-tones needed to fill the octave gap. An octave is a large leap, but it is a very consonant one and the dotted-quarter note gives plenty of time to make the leap. Allowing arpeggiation as well as passing-tones would no doubt have resulted in the transition being filled in.

The transition between the first and second motives demands some kind of interfacing, since there is an eighth-note missing from the end of the first measure (tics 6 and 7). EMOTER recognizes this problem and fills in the gap with a passing-tone:

(4.9)

```
((0 2) 50) ((2 2) 51) ((4 2) 50))
                                        ((6 2) 51)
((8 1) 53) ((9 2) 55) ((11 1) 60)
```

4.5.3 Mode Adjustment

Finally, if the phrase is in the minor mode it is checked for the possibility of needing alterations to the sixth and seventh scale degrees to conform to the Classical music style. A special operation starts at the end of the phrase and works backward to its beginning, examining each note. It changes a minor-mode leading-tone to a raised leading-tone if the next (different) pitch is higher than the present one (or the present one is the last one in the phrase). It also adjusts the sixth scale degree upward a half-step if the next pitch after it is a raised seventh. In our example there is one instance of the seventh scale degree (53) and two instances of the sixth (51). The seventh is raised a half-step (to 54) because the note following it is higher than it is. Because of this the sixth immediately preceding it is also raised a half-step (to 52). Compare Example 4.9 with Figure 4.3 to see the changes. The first sixth in the phrase, however, remains natural because the pitch following it is lower than it is.

4.6 Output of the Phrase

The backtracking control mechanism of PROLOG is such that many acceptable phrases are generated automatically. We have traced one interesting example in, it is hoped, enough detail so that the inner logic of EMOTER has been made clear. Because of time limitations and the fact that no accents, dynamic markings or other subtleties are generated by EMOTER, the phrase is output in SICCLLL notation. It is the user's responsibility (for the present) to convert the notation into standard music notation or actual sounds. In Figure 4.3 is the finished sample phrase in conventional music notation and in SICCLLL (divided by measure lines for clarity).

4.7 A Final Word

The reader should by now have a general idea how EMOTER composes (and *some* idea how a human composes, since much of the program was designed around techniques used by the author and other composers). The gradual accretion of musical concepts and thematic material from the most general emotional feelings and moods is mirrored in the sequence of steps EMOTER follows on its journey toward a finished phrase. The following chapters provide a considerably more detailed view of EMOTER's composition processes.

((-2 2) 43)	((0 2) 50)	((8 1) 54)	((16 6) 62)	((26 2) 58)
	((2 2) 51)	((9 2) 55)	((22 4) 50)	((28 2) 57)
	((4 2) 50)	((11 1) 60)		
	((6 2) 52)	((12 1) 58)		
		((13 1) 60)		
		((14 1) 62)		
		((15 1) 63)		

Figure 4.3: The Final Phrase

Chapter 5

Attributes

5.1 Introduction

Music attributes — specific properties music possesses that can help to effect moods in the music — have been discussed in Chapter 2 and mentioned in other chapters. This chapter is devoted to the closer examination of these attributes — what they are, how they are computed and how they help shape the music. A trace is included to show how the attributes are computed. The moods used as input for the trace are the moods of the main example of the thesis:

`(outgoing protesting tense active)`

As we stated in Chapter 4, the input moods must be translated into terms EMOTER can understand and use to manipulate the music. A special definition of each mood called a *moodframe* exists as a static database inside EMOTER. The moodframes for the input moods are collected, processed and combined to form a set of *attributes* that completely describe them in terms the composing routines can access, understand and use.

5.2 Obtaining Attribute Values from Moods

Table 5.1 displays the computed attributes for the above moods. The table should be carefully examined since the remainder of the chapter makes constant reference to it. (For reference information about the attributes, see Section 5.3.) We will trace the computation of only three attributes — motif, interval and motive use — since many of the attributes are determined in a fashion similar to the way motive use is computed.

The attribute values in Table 5.1 were computed independently of each other with one exception, the choice of motifs, which will be discussed below. To give the reader a feel for these values we will quickly discuss some of the more salient

Table 5.1: The Computed Attribute Values

```
((activity ((melodic 1) (harmonic 3))))
((cadences  ((half 2) (deceptive 1) (plagal 0) (authentic 0))))
((contour ((angularity 3.7) (span 6.7) (shape 0.5) (wavyness 0.6))))
((dissonance ((weak 1) (strong 0.5))))
((emotion ((value 0.8) (slope 0.5))))
((finality (0.2)))
((harmony ((naturalness 1) (chromaticness 0.5) (tonicness 0.2)
          (tempo 2.5) (angularity 1.6555))))
((intervals (((5 2) 10) ((10 7) 10) ((4 1) 7) ((8 5) 6) ((9 6) 5)
             ((7 4) 5) ((6 3) 5) ((3 0) 5) ((2 11) 4) ((1 10) 4)
             ((0 9) 3) ((11 8) 2))))
((pattern ((length 0.251) (strictness 1))))
((motiveuse (contrast 3.5)))
((motifs (
    (((43 50 51 50) minor arched
         (tense lively active sad protesting dark passionate))
      ((43 46 50) minor rising
         (tense active sad protesting outgoing passionate)))
     (((38 43 46) minor rising (outgoing passionate sad))
      ((43 46 50) minor rising
         (tense active sad protesting outgoing passionate)))
     (((51 50) minor falling
         (tense lively active longing anguished passionate sad))
      ((43 46 50) minor rising
         (tense active sad protesting outgoing passionate)))
     (((43 46 43) minor arched
         (tense static incoming dark passionate sad))
      ((43 46 50) minor rising
         (tense active sad protesting outgoing passionate))))))
((phrasing ((length 3) (naturalness 2.9661) (begin off) (end off))))
((rhythm ((activity (average 2.4494) (variance 2) (contour -1))
          (shape (syncopation (interbeat 1) (intrabeat 1))
                 (dottedness (interbeat 0.1) (intrabeat 0.1))))))
((tempo (100)))
((timesig ((2 4) 1)))

((normtimesig 2 4 8))
((measure-hierarchy (8 4 2 1)))
((metric-hierarchy (1.4142 1 0.7071 0.4999)))
```

results. Music that is outgoing, protesting, tense and active should, for example, be "active" in every musical sense of the word (emotionally, motivically, melodically, harmonically, cadentially and rhythmically). The values of these attributes are therefore quite near their maxima. Protesting, tense music is generally dissonant, with some amount of syncopation. Outgoing or protesting music generally has a sharply-rising contour at first, followed (at least in the case of protesting music) by a more gently falling contour. Since the mood is "up in the air," one would not expect much stability, as reflected in low finality and harmonic "tonicness" values. The way that the motifs and scale-degrees reflect the moods will be explained later.

5.2.1 Typical Processing Techniques

5.2.1.1 Overview

Each of the attributes uses the list of moodframes for the input moods (and in a few cases, other attributes) to compute its value. Two simple methods — averaging and voting — are used for most attributes. When the possible values are a range of (real) numbers averaging is used to obtain a single value. With a finite number of possible values, voting decides the most popular value. Table 5.2 lists the methods used to compute each of the attributes. The mode of the phrase is chosen after the motifs, all of which must be of the same mode. Motif and interval are discussed below.

Some of the moods are considered to have no bearing on some attributes. In these cases the value for the moodframe was either set to a special value any or left out entirely (see Section 5.3 for more details). For example, a passionate phrase can (in our opinion) have any number of notes per beat. A more sophisticated method for assigning values to the attributes would involve an extensive survey of the music literature. Anys are ignored in calculating the attribute values. If *all* of the input moods happen to have a value of any for some particular attribute, then a random value is chosen for that attribute.

5.2.1.2 Computing a Typical Attribute – A Sample Trace

To show exactly how a typical attribute is determined we will follow the computation of the Motive Use attribute. (This attribute determines how many different motifs the phrase should have.) First the moodframe values of motive use for the four input moods are retrieved. The values for motive use range from 0 (exact reuse of the same motif) to 4 (use a contrasting motif). The values for the tense and active moods are 3 and 4, respectively; we do not consider the other two moods (outgoing and protesting) to have any relationship on the value for motive use, so they do not appear in the moodframe database for this attribute. The numbers we do have are averaged to obtain a value of 3.5 (contrasting motifs) for motive use.

Table 5.2: Methods of Computing the Attributes

Averaging	*Voting*	*Other*
activity	cadence	intervals
contour	time signature	motifs
dissonance	phrase beginning	mode
emotion		
finality		
harmony		
motive use		
pattern		
phrasing		
rhythm		
tempo		

5.2.2 How the Interval Attributes are Processed

The scale intervals (scale degrees plus the non-diatonic pitches) are ordered by a special method, principally because unlike other attributes, we want all the intervals evaluated and compared with each other. The ordered list will be used to help choose between several candidate scale-degrees for certain embellishments. The scale intervals are all placed into a list in order of importance. More than one interval is commonly made use of, so a somewhat more sophisticated method of ordering them is used. For each interval a *discrepancy value* is calculated. Treating the list of input moods as one set and the list of moods associated with a particular interval as another set, the discrepancy value is the size of the set that is the (set) difference between the union and the intersection of the two sets. For example, if set R is {a, b, c} and set S is {b, c, d, e}, the union of the sets is {a, b, c, d, e}, their intersection is {b, c} and the set difference is {a, d, e}. The discrepancy value in this example then is three (the size of {a, d, e}). To give "importance" ratings (instead of discrepancy ratings) for the intervals, their values should vary directly (not inversely) with their importance. Therefore the computed discrepancy values for the intervals are "inverted" to make the lowest values the highest and vice versa. In addition, the tonic and dominant (perfect fifth) scale degrees are given an extra measure of importance value.

5.2.3 How the Motif Attributes are Processed

The choice of motifs for the phrase is probably the most important of the attributes. A motif is also the most complex single attribute. For these reasons, the method for choosing the motifs is the most comprehensive (and quite different from the other methods).

First of all, the major and minor motifs are processed separately. The resulting ordered lists are merged together. The mode for the phrase will be the same as the mode for the chosen list of motifs.

For each mode, every combination of motifs is examined and those whose "meanings" (combined lists of associated moods) minimally encompass the input moods are retained. Using this technique the number of motifs chosen is likely to be related to the number of moods (and therefore, to the complexity of the mood description). The combinations are ordered by their degree of agreement with the desired (computed) melodic contour. From these, the ones with the fewest members (the smallest number of *different* motifs) are chosen. If the motive use attribute calls for contrasting motives — as this example does — any single-motif solutions are removed from the list of candidates, since one motif cannot contrast with itself.

In our example, the computed contour attribute is arched (0.5). It can be built with an arched motif (liberally embellished), or two motifs — one with a rising contour and one with a falling contour. The choice made here is for a rising and an arched motif because there are no better contour choices available, as can be seen in Table 5.1 under the motifs attribute group.

5.2.4 Derived Attributes

In addition to the attributes mentioned above, three other useful facts about the time-related aspects of the phrase are computed before starting the composition phase. Unlike the others, these are derived completely from other attributes, not from the moods. They are *normtimesig*, *measure-hierarchy* and *metric-hierarchy*, and are described in the next section. The method of calculating their values is included in Chapter 6.

5.3 Representations of the Attributes

Below we define the attributes. The structure of each discussion reflects the structure of that attribute as it is implemented. At the end is an example for a particular set of moods.

TEMPO - Tempo is simply the rate of passage of the beats of the music. A range of 40 to 160 (beats per minute) encompasses the range of most tempi in actual music.

METER - Five meters were chosen as representative: 2/4, 3/4, 4/4, 6/8 and 9/8. A time signature is represented as a list of two numbers, the numerator and denominator of the time signature. For example, 2/4 is represented as (2 4). Unfortunately, this method of representation, while obvious and simple, has proven almost useless to express anything meaningful about about meter since it says nothing about the relative strengths of beats and their divisions, or about the number of divisions of a beat. Therefore, the two following representations are additionally computed and asserted after the other attribute values have been established.

NORMTIMESIG - After the beat size is calculated (for 6/8 and 9/8 the beat size depends on tempo) a list is created with three arguments: beats per measure, tics per beat and tics per measure. It is here that the relationship between tics and beats is made explicit and specific for the phrase.

MEASURE HIERARCHY - This holds the square-roots of the lengths (in tics) of each major metric subdivision level for a measure in the chosen time signature. These values are the *metric weights* of the hierarchical locations in a measure. For example, if the time signature is 6/8 with a dotted-quarter note receiving a beat and an eighth note being the smallest allowable subdivision, the MEASURE HIERARCHY list would contain (2.4494 1.7320 1). This says that a dotted-half note (6 tics) is the most important metric division with a metric weight of $\sqrt{6}$, a dotted-quarter note (3 tics) is next, with an eighth-note (1 tic) being the least important division. Although two tics divides evenly into six, it is not considered a metrical subdivision, being 2/3 of a beat.

METRIC-RHYTHMIC IMPORTANCE - At this point between the descriptions of metric and rhythmic representation we will briefly digress to discuss the method of determination of a note's metric-rhythmic "importance". This is not a static attribute like the others - asserted once before composition commences - but it describes an important property of a note nevertheless. We arrived at the method empirically, having found no theory of rhythm or meter to use. Our method seems to correlate well with intuition about rhythm and meter, however, as we hope the example below demonstrates. We wanted a representation for rhythm that would take metric location and duration into account and "return" a value for a note (independent of its pitch) proportional to its metric-rhythmic "importance" in the measure. The metric-rhythmic importance of a note is calculated as the product of its metric weight (discussed above in MEASURE HIERARCHY) and its duration (in tics).

RHYTHM - The subject of rhythm is too complex to be given a complete representation in this paper, so a few important characteristics have been chosen to represent the variety of possibilities. For more about rhythm refer to Chapter 6.

ACTIVITY - Rhythmic activity represents three aspects of the relationship of note-attacks to a beat.

AVERAGE - This is a sort of metric tempo - the average number of note attacks per beat. If a quarter-note is assumed to have the beat then the range of average notes for this program is from an eighth-note to a whole-note, which is approximately consistent with Classical music.

VARIANCE - Variance describes the durational dissimilarity of the notes in a passage. A phrase with minimum activity variance would have notes whose durations are all the same. Values range from complete *in*variance (a value of 1) to a ratio of 3:1 (expressed as a value of 3) between the average duration and the maximum deviation (either shorter or longer than average). A typical result of activity variance (without syncopation) in music is the use of so-called dotted rhythm (a long note on the beat followed by a shorter note just before the next beat). The passage below has an average activity of two tics, with a variance of two. Only the duration of the notes is given.

```
( ((* 2) *) ((* 1) *) ((* 1) *) ((* 4) *) )
```

SYNCOPATION - Simply put, a syncopated note lasts through a stronger beat or beat division than it started on. Another way of defining syncopation compares the actual location of a note with its expected location. Most syncopated notes occur just before they normally would.suggests that at least two levels of syncopation exist. While this definition fits much better with Meyer's theories, we found that the definition stated above is both essentially equivalent and easier to implement. (Another level exists - *metric* syncopation - but for simplicity's sake this was not included.) *Interbeat* syncopation describes the fraction of *beats* that are syncopated, while *intrabeat* syncopation describes the fraction of notes that are *non-beat* syncopated notes (i.e., start off the beat). While the range of possible fractional syncopation values is between zero and one, an unaccompanied phrase whose notes are *all* syncopated would have no beat-based foundation, and thus would not sound syncopated! Therefore, the largest fraction of syncopated notes should not be more than about half the total.

DOTTEDNESS - Like syncopation, dottedness comes in two flavors: interbeat and intrabeat. Interbeat dottedness is the ratio of dotted rhythms spanning more than one beat to the total number of spans of the same length. Intrabeat dottedness is the ratio of dotted beats to the total number of beats.

DISSONANCE - Since this experiment does not allow or need the existence of simultaneous tones (chords), the definition of *dissonance* has been simplified to apply only to single tones. In the context of Classical music a dissonant tone is simply one that is not a member of the current (implied) harmony. The measure of dissonance is the fractional duration of the non-chord tones. A value of zero would mean that all pitches are chord-tones. Like dissonance, if all of the pitches were non-chord tones, they would seem to define a harmony themselves and lose their sense of dissonance. So the largest fractional value of dissonant tone durations should be something less than one-half. Also like syncopation, dissonance has two levels based again on rhythm and meter. *Weak* dissonance is *unstressed* — occurring on a weaker metric location than the consonance (chord-tone) associated with it. Weak dissonances are passing-tones, grace-notes and neighboring-tones. *Strong* dissonances occur on stronger locations than their resolutions. The two strong dissonances are suspensions and appogiaturas.

HARMONY - A chord in the implicit chord-structure of a note-group is represented as a triplet of integers: the relative (normalized diatonic) tonic, third and fifth of that chord. For example, the dominant (V) chord in either mode is (5 7 2).

HARMONIC MOTION - Two aspects of the way harmony changes are represented here.

TEMPO - The number of chord changes in a phrase is called the tempo of the harmonic motion. Most phrases have from zero to eight chord changes.

ANGULARITY - The number of common tones between chords is a convenient way of determining the amount of harmonic motion, or angularity. Assuming only triadic (three-pitch) chords, angularity can range from no change to all three pitches changing.

CADENCE TYPE (harmonic finality) - The chord or chord progression with which a phrase ends greatly affects the impression of finality of that phrase. Some common cadences, from least to most final, are:

- half, which explicitly leaves the listener "up in the air";
- deceptive, which deceives the listener into believing the cadence is final until the last chord;
- plagal, the "amen" church chorale cadence;
- imperfect authentic, a final cadence in which the last melodic pitch is the third or fifth; and
- perfect authentic, the same final cadence as above but with the last melodic pitch being the tonic. This cadence is normally reserved for the end of the melody since it is the most final one.

PITCH - the following attributes describe aspects of pitch that are independent of time and harmony.

IMPORTANT SCALE INTERVALS relative to the tonic - Certain scale degrees are especially characteristic for certain moods, according to [Cooke 83] These can be any subset of the members of the diatonic scale.

DISJUNCTNESS - Disjunctness or melodic angularity is the average of the absolute values of the intervals between adjacent pitches. This attribute captures the degree of smoothness of a melodic passage. A disjunctness value of zero would indicate a melody of one pitch, a value of a third or fourth would be more common, and a disjunctness of an octave or so would be found in only the most angular 12-tone pieces.

SCALE ACTIVENESS - The ratio of active-note (2nd, 4th, 6th and 7th scale degree) durations to total duration in the phrase is its scale activeness. These scale degrees have tendencies to "move" to the tonic, third or fifth, so their presence tends to make a phrase feel more "mobile." Values of activeness range from zero (actually almost impossible since, in the absence of the above active tones, the third and fifth seem active) to one hundred percent.

MODE - In Classical music two scales dominated. They are called the major and minor modes. [Cooke 83], [Hevner 35] and others have shown that these modes profoundly affect the mood of the music. They are represented using the same scales, but with a different number as tonic.

CONTOUR - The "physical" location of the notes in a phrase as it "moves" through time is its contour. Three aspects of contour are given here.

SHAPE - The rising and falling "motions" of a phrase can be very roughly categorized into four common shapes: rising, arched, level, and falling. The importance of this aspect of contour is, according to experiments done by [Hevner 35] of minimal importance in establishing any particular mood, although the highest point of a passage usually coincides with its climax.

SPAN - Span is the difference between a phrase's highest and lowest pitches. More dramatic phrases seem to have wider spans. A two-octave maximum was deemed reasonable here.

MELODIC FINALITY - The activeness of the last note in a phrase contributes to its sense of finality almost as much as the kind of cadence. Regardless of the cadence, the tonic of the scale feels the most final, followed by the third or fifth. Any other tone will give little or no feeling of finality.

PHRASING The phrase as a whole has certain time-related attributes that seem worth including.

LENGTH - According to [Lundin 53] the average phrase length varies from about two to eight seconds, with four or five being the average. No relationship between phrase length and mood was found documented in the literature, however. Since some measure of actual time-span was necessary, some rough estimates were made for a few moods and random choices centered about five seconds are generated by EMOTER for the rest.

NATURALNESS - Most phrases are a small power of two measures long. Therefore, the most "natural" phrase is exactly this length, with less natural ones up to one-third longer or shorter than the ideal.

FINALITY - The metric strength of the last note affects the character of a phrase enough so that the names "feminine ending" and "masculine ending" [Piston 62] have been traditionally used for phrases ending on a weak or strong beat, respectively.

BEGINNING - The metric location of the *first* note of a phrase also affects the feel of the phrase. The possible choices have been simplified to having the phrase either start or not start on the downbeat. Decisions here were made based on our own experience.

EMOTION - The word "emotion" is used here in a slightly more general way than [Meyer 56] uses it. Nevertheless it is a useful attribute, helping to make decisions of degree. Many moods have emotional connotations, so it was included as an attribute. Two aspects of EMOTION were singled out.

LEVEL - This is the average "emotional" content of the passage. No units of emotional quantity could be found, so an arbitrary unitless scale of zero (minimal emotion) to one (maximum emotional level) was set up.

SLOPE - Emotion can increase or decrease (ignoring emotional peaks at climaxes), so SLOPE specifies this aspect. Again no units were available so a scale of minus one (maximally decreasing) to plus one (maximally increasing) was chosen.

MOTIVE USE - A motive can be treated in a number of basic ways: it can be *reused*, either by *exact* or *varied* repetition or sequence (starting on a new pitch), or a *contrasting* motive may replace it temporarily (by definition, a motive must be used at least twice to qualify as a motive, so in the latter case the original motive must return eventually).

PATTERN - A *pattern* is a group of notes having certain properties which are capable of being imitated. Aspects of a pattern that are implemented as attributes are:

STRICTNESS of imitation - This has the same meaning and representation as for motive use above, but this has a wider application.

Fractional LENGTH of a pattern member - In a phrase a pattern can be comprised of one to four members for the purposes of this project. A single member in a phrase indicates that pattern imitation occurs among phrases instead of within.

MOTIFS - [Cooke 83] listed about ten basic motifs which he found over and over again occurring in Classical melodies. A motif is represented as a four-tuple consisting of its normalized pitches (chromatic, centered around middle C or the A below it depending on mode), its contour, mode and "meaning" (a list of moods associated with the motif).

5.4 Conclusion

Let us close this chapter with two caveats concerning the attributes. While we have no doubt that the attributes discussed here do affect the perceived mood of music they control, these attributes are not the only ones operating on music, even on the mere abstracted melodic material dealt with here. Furthermore, the attributes chosen for this thesis have greater, deeper and more varied applications than initially anticipated. They are a very powerful expressive tool in music composition, and we have barely started to realize their full potential. Nevertheless their performance in EMOTER is convincing evidence, we believe, of the tremendous utility and complexity of the attributes of music.

Chapter 6

A Theory of Rhythm

6.1 Introduction

Judging from the small number of books about rhythm, it is not well understood [Cooper & Meyer 60], [Yeston 76]. The sources we read discussed rhythm from a purely syntactic perspective. As with most music theory texts, the "rhythm theorists" seem to be only interested in labelling rhythms. Nothing is said about the reason why the composer chose a particular rhythm. While these books are undoubtedly useful for analytical labeling of music, they do not shed any light on theories behind the composition of rhythms appropriate to the mood and emotion of the music, or even the relationship between pitch and rhythm. Cooke's theories include somewhat more useful ideas on the relationships of mood and rhythm, but they are much too general to offer any help when specific choices must be made. [Meyer 56] discusses rhythm only in relation to other rhythms and how a *change* of rhythm affects the emotional content. While somewhat helpful, his insights do not prescribe a theory of "rhythmic composition." We therefore have over the course of this research found it necessary to develop an original *generative* theory of rhythm and its relationship to pitch. Without this theory EMOTER would have no "idea" how to assign durations to the pitches it chooses. Although a small amount of work is needed to complete the theory, it has so far shown itself to be extremely useful and has produced rhythms that appropriately complement the melodic material.

According to the theory, one of the principal purposes of rhythm and meter is to clarify and reinforce the hierarchy of melodic pitch-functions — basically Schenker's background-middleground-foreground pitch hierarchy [Schenker 56]. Rhythm lends emphasis by setting up a hierarchy of *durations*. Meter does so by setting up a hierarchy of *cyclic locations* in time. By this we mean a regularly recurring rhythmic pattern, set up as a background expectation. The period of recurrence is made clear by giving some parts of the pattern more "importance" than others. Initially rhythm, tonality and meter can be established if they all support one another. We are better able to recognize the more important pitches (e.g., chord-tones and important scale-degrees) when they are highlighted *a*) *metrically* by placing them on "stronger" locations in the measure than the less important pitches, and/or *b*) *rhythmically* by giving them longer durations. Another highlighting technique, *syncopation*, will be discussed later in the chapter.

For metrical highlighting to be effective, the meter must first be established. To establish a meter as the norm, pitch, rhythmic duration and metric location must all "cooperate" initially, so that important pitches or long notes fall on strong metric locations. Thereafter (in Classical music) the meter becomes one of the global norms (along with tempo, mode, key signature and the like), barring any extreme syncopation.

6.2 Meter

A time signature is initially represented in EMOTER as a list of two numbers, the numerator (traditionally the number of beats in a measure) and denominator (the kind of note that gets a beat) of the time signature. For example, $\frac{2}{4}$ is represented as (2 4). Unfortunately, this method of representation, while obvious and simple, has proven almost useless to express anything meaningful about meter since it says nothing about the relative strengths of beats and their divisions, or about the number of divisions of a beat. Therefore, the two following representations are additionally computed and added to the information EMOTER can draw upon.

6.2.1 Normalized Time Signature

After the beat size is calculated (for $\frac{6}{8}$ and $\frac{9}{8}$ the beat size depends on tempo) a list is created with three arguments: beats per measure, tics per beat and tics per measure. (Recall from the representation chapter that a tic is the smallest duration in a particular phrase.) It is here that the relationship between tics and beats is made explicit and specific for the phrase.

6.2.2 Metric Hierarchy

It is not enough to know how many tics are in a beat, since the first beat of a measure generally carries more weight than the second. We therefore need a finer-grained hierarchy to represent the meaning of meter — the relative strengths of the measure- and beat-divisions. After some experimentation (and based on our knowledge of music) we decided upon the following hierarchy for a measure of duple meter (an even number of beats): the first beat is the most important, followed by the "middle" beat (e.g. the third beat in $\frac{4}{4}$), followed by the others. In triple meter the third beat is considered second in importance to the downbeat.

Fractions of a beat have the same hierarchy — another example of self-similarity in music.

Furthermore, relative numeric values called *metric weights* have been assigned for each of the hierarchical beats and beat-fractions. The method of determination of the values was empirical — we simply examined a large number of rhythms and found a mathemetical relationship that seemed to represent their inner structures adequately. The weights we chose are the square-roots of the lengths (in tics) of each major metric subdivision level for a measure in the chosen time signature. For example, if the time signature is $\frac{6}{8}$ with a dotted-quarter note receiving a beat and an eighth note being the smallest allowable subdivision, the measure-hierarchy list would contain (2.4494 1.7320 1). This says that a dotted-half note (6 tics) is the most important metric division with a metric weight of $\sqrt{6}$, a dotted-quarter note (3 tics) is next, with an eighth-note (1 tic) being the least important division. Although two tics divides evenly into six, it is not considered a metrical subdivision, being $\frac{2}{3}$ of a beat (in $\frac{6}{8}$ time).

Metric weighting helps to determine a note's metric-rhythmic importance. We wanted a representation for meter and rhythm that would take metric location and duration into account and "return" a value for a note (independent of its pitch) proportional to its metric-rhythmic "importance" in the measure. The metric-rhythmic importance of a note is calculated as the product of its metric weight and its duration (in tics). Within our work this weighting found application in determining the rhythm of a skeletal motive and in the rhythmic embellishment of the motive.[1].

6.2.3 Syncopation

Simply put, a syncopated note lasts through a stronger beat or beat division than it started on. Another way of defining syncopation compares the actual location of a note with its expected location. While this second definition fits much better with Meyer's emotion theories because it relates syncopated rhythms to expectation and frustration, we found that the first definition is both essentially equivalent and easier to implement. (Another level — *metric* syncopation — exists, but for simplicity's sake this was not included.) The opposite of syncopation is (metric) stability. Even and dotted rhythms are stable, with dotted rhythms being more so because they emphasize the stronger metric location (with a longer duration). Stable rhythms are generally used when the finality attribute has a high value.

Interbeat syncopation describes the fraction of *beats* that are syncopated, while *intrabeat* syncopation describes the fraction of notes that are *non-beat* syncopated

[1]Unfortunately, we had to abandon this useful representation in the case of skeletal rhythm determination. The method that made use of metric-rhythmic importance was a generate-and-test strategy. The strategy was so computationally expensive that running it often overflowed the internal stack in our version of PROLOG.

notes (i.e., start off the beat). While the range of possible fractional syncopation values is between zero and one, an unaccompanied phrase whose notes are *all* syncopated would have no beat-based foundation, and thus would not sound syncopated! Therefore, the largest fraction of syncopated notes should be less than about half the total.

6.3 Application of the Theory

6.3.1 Introduction

In this section we discuss two places in EMOTER where the rhythmic theory is put to work. First we examine the process whereby a foundational rhythm is assigned to the rhythm-less motif. Included is a trace of the skeletal rhythmicization of a motif. Following that we will describe how the various embellishments help determine rhythm.

6.3.2 Assignment of a Skeletal Rhythm

A metrically stable skeletal rhythm is computed, based on our theory, for the first occurrence of every motif. To do this, the relative importances of the motif pitches and the relative metrical strengths of the possible locations for the pitches must be calculated. We briefly digress from the subject of this chapter to mention pitch.

6.3.2.1 Pitch Importance

Pitch importance is taken from the interval attribute. Recall from Chapter 5[2] that the fifth scale degree (a pitch of 50) was determined to be the most important. The weighted list for the first motif of our example phrase is:

```
((43 7) (50 10) (51 8) (50 10))
```

6.3.2.2 Calculation of the Metrical Contour

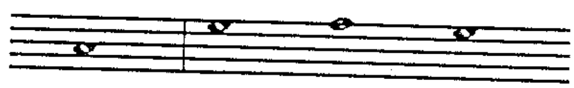
Metrical contour is the metrical "shape" of a measure, the metrical strengths of locations within it where a note is allowed to exist. To calculate the metrical strengths of the possible locations it is first necessary to determine what locations are "possible." The number of locations should be at least twice the number of motif pitches, to allow the notes ample space for different arrangements; the locations themselves should be equally spaced throughout the time-slot allocated for the motive.

[2]See Section 5.2.2 and the intervals list of Table 5.1.

The locations must be represented in such a way as to allow "joining" some of them with the motif pitches. For this purpose only, a representation visually similar to SICCLLL, called a *weighted-locations template* is used to represent the possible note locations (which we will call *slots*) and their metric strengths within the span of a motive-length. Each slot of the template has the form
((<pitch> <pitch-weight>) <metric-weight>).
<Pitch> and <pitch-weight> are initially set to zero, indicating no note present; <metric-weight> of course has the metric weighting for the slot. The metric hierarchy attribute supplies the metric weightings. In this example the time signature is $\frac{2}{4}$ and there are four pitches in the first motif. Therefore, eight "notes" are needed for the template, with each note having a duration of a sixteenth-note. The weighted-locations template looks like this before the pitches are inserted:

```
(((0 0) 1.4142) ((0 0) 0.5) ((0 0) 0.7071) ((0 0) 0.5)
((0 0) 1.0) ((0 0) 0.5) ((0 0) 0.7071) ((0 0) 0.5))
```

In each "note" of the motif-template the first zero is replaced by the motif pitch and the second zero is replaced by the interval-weighting of the pitch. An attribute of phrasing determines whether the passage is to start on or off (before) the first beat. If the choice is off the beat, the metrically-strongest starting location for the first motive is found before any other time-locations are assigned. This is done by "rotating" the pitches of the motif and finding the best rhythmicization for each rotation (using the template). To keep track of the true beginning of the motif, its first pitch is made negative. Since in our example the fifth scale degree is the most important, EMOTER decided that it should be placed at the most important metrical position — the downbeat. The rotation chosen was therefore:



(6.1)

Specific metrical locations must be found for the rest of the motif pitches. Briefly, the most important pitch is given the metrically strongest location, the second-most important pitch is given the second-most important location, and so on. More precisely, to find the best metrical placement in the template for them, the following method is used. Place the most important pitch in the metrically strongest slot. (The fifth was placed in the downbeat slot.) Split the motif and the template at that pitch and slot, and recursively find the best metrical placement in each template for each fragment. If there is no available location for a pitch, then backup, choose the second-most important location for the most important pitch instead and try that arrangement. This method returns the most appropriate metrical location possible for each pitch according to its importance. The filled template ("rotated") and its musical representation (here and from now on transposed up an octave for clarity) are shown below.

(6.2)

```
(((43 7) 0.7071) ((0 0) 0.5) ((50 10) 1.4142) ((0 0) 0.5)
((51 8) 0.7071) ((0 0) 0.5) ((51 10) 1.0) ((0 0) 0.5))
```

6.3.3 Rhythm and Embellishment

We briefly discuss rhythmic aspects of each of the embellishment implementation routines. As stated in Chapter 4, the principal decision-making these routines do is rhythmic — the pitch has already been chosen (and a general rhythmic placement also, possibly). Since the skeletal rhythm assigned to a motif is purposefully "stable", it is also rather dull. If a motive is to have an interesting rhythm, therefore, it will be given by one or more embellishments to the skeletal motive. One byproduct of embellishment is an increase in rhythmic activity — the density of notes in a given time-span. Since rhythmic activity is one of the attributes, it is used to control the amount of embellishment done to a motive.

Appogiatura - The appogiatura includes the choice of interbeat or intrabeat syncopation. If the former is chosen, both the appogiatura and its resolvent must fall on beats. Otherwise the resolvent should fall off the beat.

Escape-tone - The escape-tone is handled similarly to the appogiatura except that the choice is between interbeat and intrabeat *dottedness*, or possibly an even rhythm if there is no choice. The escape-tone must fall on a rhythmically "stable" location in any event — no hint of syncopation is allowed here. The technique for insuring this is to find the metrically strongest location just before the second note of the note-pair, such that the location is at least as far from the first note as it is from the second.

Neighboring-tones - Depending on whether the neighboring-tones are to occur before or after the original note, the *consonant* member of the neighboring-tone pair is treated rhythmically as an escape-tone or an appogiatura, respectively. Of course, room for the dissonant tone must be allocated between the two notes. It is then sandwiched between the two (identical) consonant pitches as evenly as possible (becoming a dotted rhythm only when necessary).

Arpeggiation, Anticipation, Repeated Notes - These embellishments basically use the same routines as the appogiatura and escape-tones, as they are rhythmically similar to them.

Passing-tones - The passing-tones are rhythmically placed as evenly as possible. If there is not exactly the correct amount of room for them, more duration will be given to the first ones. For example, if two passing-tones had to fit between the second and third beat of a measure of $\frac{4}{4}$, they would be placed on the third and fourth sixteenth-notes of the second beat.

6.4 Conclusion

The rhythmic-metric theory stated above has been implemented in EMOTER, with highly satisfactory results. At this point, the rhythms generated may be the most interesting aspect of the phrases. Even more exciting to us is further development of this theory. We see implications of the theory in metric and "weak" syncopation, more sophisticated rhythmic elaboration of motives, implementation of rhythmic contour, more "interesting" phrase-lengths and a better relationship with Meyer's grammar, to name just a few.

Chapter 7

Embellishment and Elaboration

7.1 Introduction

In this chapter we concentrate on the two methods by which EMOTER develops simple musical entities into more interesting, sophisticated ones. The motivation behind embellishment and elaboration was discussed in Chapters 2 and 4. Here we will show how and where they function. The trace running through the chapter is the trace of our main example. The final sections define the embellishments and elaborations themselves.

7.2 Embellishment of the Motive

The skeletal motive is embellished by appropriate insertion of auxiliary notes into the passage until it is determined that activity, dissonance and angularity requirements are at least approximated. The choice of which notes are to be embellished is determined by evaluating all "correct" possibilities, choosing the best one, re-evaluating the embellished motive to take the new addition into account, and repeating the process until the above requirements tell it to stop. At the end of this section we will trace the highlights of the progress of skeletal motive 2 as it is embellished. Skeletal motive 2 is shown[1] in Example 7.1.

7.2.1 Evaluating the Need for Embellishment

A motive must not be over-embellished. Since embellishment consists entirely in adding notes to a motive by splitting up ones already present, every embellishment will increase the activity of the motive. The process of embellishment is halted when the motive's rhythmic activity exceeds the value of the activity attribute. At that time the motive is ready to be incorporated into the phrase. This evaluation

[1]The last note displayed is not part of the motive. Its function will be explained later.

is performed each time an embellishing note or note-group is added to the motive. Our rhythmic activity attribute value is 2.4494 attacks per beat, so we should expect about nine or ten notes in the embellished four-beat motive. By the way, it is possible that no embellishment will be performed on a motive even if the evaluation succeeds. This would occur if none of the *individual* notes in the motive were embellishable (for example, if there were no room for more notes or the pitches were not consonant).

7.2.2 Choosing an Embellishment

Choosing the best embellishment for a motive consists in first computing and evaluating all candidates for each note of the motive, then choosing the best embellishment and checking whether implementing it will violate any attributes (such as *maximum* rhythmic activity) that are "global" to the motive. To do this each adjacent note-pair in the motive is examined for eligibility of embellishment between the two notes. The evaluation is based on the maximum number of embellishing notes allowed for the pair (how many tics long the first note is) and which note(s) is/are embellishable. An embellishable note must both be a chord-tone or an important scale degree and have an acceptible metric strength (how acceptible is determined individually by each individual embellishment rule). Any note-pair may have passing-tones inserted even if it is not embellishable, however, as long as there is room for them (horizontally *and* vertically). The "dummy" last note is not embellishable. Below we briefly discuss the requirements for each type of embellishment to be considered as a possible candidate for embellishing a note-pair. These requirements are implemented inside separate routines for each embellishment. Unlike previous examples, this one is iterative and will be included in Tables 7.1 through 7.5. For definitions of the embellishments see Section 7.4.

Appogiatura -

- The first note of the note-pair must be on a strong beat.
- Syncopation must be allowed by the syncopation attribute (the most appropriate type — between interbeat and intrabeat — is chosen).
- Strong dissonance (dissonance occuring on a beat) must also be allowed by the dissonance attribute.
- The first pitch must be consonant.
- The presence of the appogiatura should not worsen the match between the melodic angularity attribute value and the angularity of the note-pair. The routine chooses between the upper and lower appogiatura for the best match.
- The melodic activity attribute should be greater than zero.

Escape-tone -

- The escape-tone should be an important interval.
- The second pitch of the note-pair must be a chord-tone.
- Dissonance at a weak metric location must be allowed by the dissonance attribute.
- The second note of the note-pair must occur on a beat.
- The presence of the escape-tone should not worsen the match between the melodic angularity attribute value and the angularity of the note-pair. The routine chooses between the upper and lower escape-tone for the best match.
- In addition to evaluation of the appropriateness of an escape tone, the routine computes its pitch: if the note-pair is rising, the escape-tone is a step above the second pitch; otherwise it is a step below the second pitch.

Neighboring-tones - The neighboring-tone figure can occur before or after a note (or both), so both of these possibilities must be explored. In either case:

- The note to be given the embellishment must occur on a strong beat.
- The melodic angularity must be smooth
- There must be room for at least *two* (extra) notes in the duration of the first note.
- Weak dissonance must be allowed by the dissonance attribute.
- The dissonant tone should be an important scale degree. It may be just above or below the resolvent tone; both are evaluated and the better one is chosen to represent the neighboring-tone.
- If the finality is high, the rhythm dottedness attribute should allow a dotted rhythm (if it is embellishment of the second note of the pair that is being evaluated), which is considered to be more "final" than a syncopated or even rhythm.

Anticipation or Repeated Note -

- This embellishment and the next one are consonant. Therefore, the dissonance attribute values must be low for both if either is to be applicable.
- The melodic angularity must be very smooth since a repeated pitch is being considered.

- If the finality is high, the rhythm dottedness attribute should allow a dotted rhythm (if it is embellishment of the second note of the pair that is being evaluated), which is considered to be more "final" than a syncopated or even rhythm.

Arpeggiation -

- Fit the melodic angularity to the arpeggiation: if desired angularity $\leq$ angularity between the note-pair, then find an arpeggiation note between the pair or give up; otherwise, try all chord tones above and below until the desired angularity is exceeded and choose the best "fit".
- No wide fast leaps should result from this embellishment.
- If the finality is high, the rhythm dottedness attribute should allow a dotted rhythm (if it is embellishment of the second note of the pair that is being evaluated), which is considered to be more "final" than a syncopated or even rhythm.

Passing-tones While not really an embellishment since it "connects" two notes, a set of passing-tones is treated as one for convenience.

- There must be at least enough room between the members of the note-pair for the correct number of passing tones to join them.
- A smooth contour should be called for.
- Metrically weak dissonance must be permitted by the attribute value.
- The greater the fraction of passing tones that are important intervals, the better.
- The second note must not be the "dummy" last note.

All of the evaluated embellishment candidates (for the entire motive) chosen by the above routines are kept in an ordered list, best embellishment first.

7.2.3 A Pre-Embellishment Implementation Check

Before an embellishment is chosen from the above list to be added to the motive there is one final check. Normally, the first embellishment on the list is chosen to be implemented, but first it is given a test to be sure that it will not make the motive too rhythmically "active". The maximum number of attacks per beat is calculated from the rhythm attributes and compared with the number of notes that will be appearing in the newly-embellished motive. If the maximum number is exceeded, the next embellishment on the list is chosen and tested. If every embellishment would cause the motive to become too active, the embellishment process is stopped.

7.2.4 Implementation of an Embellishment

Assuming that the pre-embellishment-implementation check allows an embellishment to be chosen, the embellishment must be "installed" into its assigned location within the motive. As a byproduct of the evaluation process above, the embellishing pitch(es) for all of the embellishments have already been chosen. Implementation of the "best" embellishment is therefore fairly straightforward. If there is a choice of locations some rhythmic decisions will have to be made. Even here some of the embellishments carry some rhythmic information with them as indicated in their requirements. After that, insertion of the embellishing notes is trivial. Chapter 6 discusses the rhythmic aspects of the embellishment implementation routines.

7.2.5 Continuation of Embellishment

After implementing the chosen embellishment but before going on, all embellishments associated with the note-pair that was embellished are discarded. They are probably useless now because the former note-pair is now a note-triplet (or quadruplet, or whatever) and has different characteristics from the old note-pair. However, there still may be room for embellishment among the new notes. If the motive as a whole still needs more embellishment (see **Evaluating the Need for Embellishment** above), only this new section will be tested for allowable embellishments; the old evaluations for note-pairs not touched are still valid. This process of recurring embellishment will continue (recursively) until the motive has been embellished enough.

7.2.6 A Trace of the Embellishments Made to Motive 2

After the dummy note has been added to skeletal motive **2** it looks like this:

(7.1)

```
(((8 4) 43) ((12 4) 46) ((16 6) 50) ((22 0) 43))
```

There are three note-pairs to be evaluated for embellishments. Tables 7.1 through 7.5 display a summary of the results.

Table 7.1: Sample Trace of Embellishments, Part 1

Embellishments initially proposed, listed by location:

```
(((8 4) 43) ((12 4) 46))
     ((passing (45)) 0.6646)
     (((appogiatura intrabeat) (41)) 0.7257)
      ((escape (48)) 0.7613)
      (((neighbor pre) (46 48)) 0.6148)
      (((neighbor post) (41 43)) 0.6286)
      (((repeat post) (43)) 0.4907)
      (((repeat pre) (46)) 0.2314)

(((12 4) 46) ((16 6) 50))
     ((passing (48)) 0.7313))
     ((escape (51)) 0.7821)
     (((neighbor pre) (50 51)) 0.7398)
     (((neighbor post) (48 46)) 0.5870)
     (((repeat post) (46)) 0.4907)
     (((repeat pre) (50)) 0.2314)

(((16 6) 50) ((22 0) 43))
     (((arpeggiation post) (46)) 0.5690)
```

Embellishment chosen and results:

```
escape (51) between ((12 4) 46)|     and     |((16 6) 50)
                                |            |
      (((8 4) 43) ((12 3) 46)|((15 1) 51)|((16 6) 50) ((22 0) 43)))
```

(7.2)

Table 7.2: Sample Trace of Embellishments, Part 2

Embellishments "left over", but still eligible as future embellishments:

```
(((8 4) 43) ((12 3) 46))
    ((passing (4 5)) 0.6646)
    (((appogiatura intrabeat) (41)) 0.7257)
    ((escape (48)) 0.7613)
    (((neighbor pre) (46 48)) 0.6148)
    (((neighbor post) (41 43)) 0.6286)
    (((repeat post) (43)) 0.4907)
    (((repeat pre) (46)) 0.2314)
(((16 6) 50) ((22 0) 43))
    (((arpeggiation post) (46)) 0.5690)
```

New embellishment candidates added as a result of the previous embellishment:

```
(((12 3) 46) ((15 1) 51))
    ((passing (48 50)) 0.7270)
    (((neighbor post) (48 46)) 0.5870)
    (((repeat post) (46)) 0.4907)
(((15 1) 51) ((16 6) 50)): no embellishments possible here because
    there's no room between the notes
```

Embellishment chosen and results:

```
escape (48) between ((8 4) 43)|     and     |((12 3) 46)
                                |           |
                  (((8 3) 43)|((11 1) 48)|((12 3) 46) ((15 1) 51)
                   ((16 6) 50) ((22 0) 43))
```

(7.3)

Table 7.3: Sample Trace of Embellishments, Part 3

Embellishments left over, but still eligible:

```
(((12 3) 46) ((15 1) 51))
    ((passing (48 50)) 0.7270)
    (((neighbor post) (48 46)) 0.5870)
    (((repeat post) (46)) 0.4907)))
(((16 6) 50) ((22 0) 43))
    (((arpeggiation post) (46)) 0.5690))))
```

New embellishment candidates added as a result of the previous embellishment:

```
(((8 3) 43) ((11 1) 48))
    ((passing (45 46)) 0.6270)
    (((appogiatura intrabeat) (41)) 0.7530)
    (((neighbor post) (41 43)) 0.6286)
    (((repeat post) (43)) 0.4907)
(((11 1) 48) ((12 3) 46)): no embellishments possible here
```

Embellishment chosen and results:

```
appogiatura (41) between ((8 3) 43)|   and    |((11 1) 48)
                                   |          |
                        (((8 1) 41)|((9 2) 43)|((11 1) 48)
                         ((12 3) 46) ((15 1) 51) ((16 6) 50)
                         ((22 0) 43))
```

(7.4)

Table 7.4: Sample Trace of Embellishments, Part 4

Embellishments left over, but still eligible:

```
(((12 3) 46) ((15 1) 51))
    ((passing (48 50)) 0.7270)
    (((neighbor post) (48 46)) 0.5870)
    (((repeat post) (46)) 0.4907)
(((16 6) 50) ((22 0) 43))
    ((((arpeggiation post) (46)) 0.5690)
```

New embellishment candidates added as a result of the previous embellishment:
none
Embellishment chosen and results:

```
passing (48 50) between ((12 3) 46)|          and          |((15 1) 51)
                                   |                       |
  (((8 1) 41) ((9 2) 43)           |                       |
           ((11 1) 48) ((12 1) 46)|((13 1) 48) ((14 1) 50)|((15 1) 51)
           ((16 6) 50) ((22 0) 43))
```

(7.5)

Table 7.5: Sample Trace of Embellishments, Part 5

Embellishments left over, but still eligible:

```
(((9 2) 43) ((11 1) 48))
    (((repeat post) (43)) 0.4907)
(((16 6) 50) ((22 0) 43))
    (((arpeggiation post) (46)) 0.5690)
```

At this time EMOTER decides that motive 2 has been embellished enough. After the dummy last note has been removed the final version of the motive is:

```
(((8 1) 41) ((9 2) 43) ((11 1) 48)
((12 1) 46) ((13 1) 48) ((14 1) 50) ((15 1) 51)
((16 6) 50))
```

The motive can only now be transposed to its assigned pitch-area. After transposition (up an octave) the motive is:

(7.6)

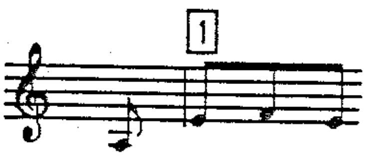
```
(((8 1) 53) ((9 2) 55) ((11 1) 60)
((12 1) 58) ((13 1) 60) ((14 1) 62) ((15 1) 63)
((16 6) 62))
```

7.3 Elaboration

In our example, the third phrase-member is to be a FRUSTRATEd elaboration of motive 1 (see Example 7.7). The choices of elaborations to FRUSTRATE a motive are, approximately from least to most emotional (in our opinion):

1. reorder pitches (any two)
2. cadence (half or deceptive)
3. reembellish the skeletal motive
4. remove a consonant note
5. generate a new skeletal motive (and embellish it)

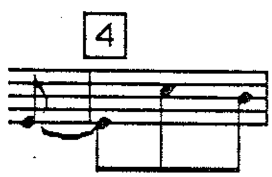
The phrase being composed is to be quite emotional. One of the more emotional operations, removing an important note (creating strong expectation for its return) was therefore chosen by EMOTER to operate upon motive 1. In particular, it chose the most important single note in the motive to remove, the fifth scale degree occurring on the downbeat. In such a case the "hole" is filled in by extending the previous note by an amount equal to the duration of the removed note. (The hole could have been left there — as a rest — but we felt that rests were not required to demonstrate the thesis.) Here is the original motive:



(7.7)

```
(((-2 2) 43) ((0 2) 50) ((2 2) 51) ((4 2) 50))
```

and the FRUSTRATEd motive, transposed (finally):



(7.8)

```
(((22 4) 50) ((26 2) 58) ((28 2) 57))
```

7.4 Definitions of the Embellishments

Below we define the embellishments used by EMOTER. The small musical examples to the right of the embellishment names are intended to help the reader who is unfamiliar with these musical devices "visualize" them more clearly. The open notes represent the embellishment and the solid note represents the note that is embellished. The vertical bar places the note following it on a beat, as if that note were the first in a measure.

PASSING-TONE - The PASSING-TONE is the only embellishment that does

not embellish a particular motive note. Rather, it fills in scale steps between two adjacent notes (possibly between two adjacent motives). It has one argument, a list of the fill-in pitches (without any rhythm).

NEIGHBORING-TONE - The NEIGHBORING-TONE embellishment can either occur before or after a motive note, and it can be an upper- or lower-neighboring-tone (or both). Therefore its representation has two arguments, one for upper/lower/both and one for before/after the note.

APPOGIATURA - The APPOGIATURA is the only accented dissonant embellishment we use. Since it (and other accented embellishments) rhythmically displaces the note it embellishes, it has an argument to describe approximately how far to displace (within a beat or a multiple of a beat). There is also an argument that indicates whether an upper- or lower- appogiatura is meant.

ESCAPE-TONE - An ESCAPE-TONE is similar to an APPOGIATURA except that it is unaccented. Thus the ESCAPE-TONE has the same arguments as the APPOGIATURA.

ARPEGGIATION, REPEAT - While the preceding embellishments have been

built on non-chord tones, the ARPEGGIATION and the REPEAT embellish by reinforcing the current harmony (the repeat also reinforces the melody-tone). Both have an argument similar to NEIGHBORING-TONE for before/after the note. In addition, ARPEGGIATION has an argument to hold the chosen chord-tone.

7.5 Definitions of the Elaborations

REPEAT - This REPEAT repeats the entire passage, not just one note. No arguments are necessary.

TRANSPOSE - TRANSPOSE moves the entire passage up or down one of the intervals in its first argument (randomly chosen), keeping it in the same key. This also changes the harmony of the passage, however, so TRANSPOSE has a second argument specifying whether the harmony is to be adjusted back to the original harmony of the passage (a deceptively tricky task due to the presence of different intervals in a chord) or left as is. TRANSPOSE is the elaboration used to construct sequences.

REORDER PITCHES - This elaboration exchanges two somewhat randomly selected pitches with each other. One argument allows some choice: either or both pitches can be chord-tones or not.

REMOVE NOTE - This deletes a note (an argument allows selection of consonant or dissonant) and fills the "hole" by adding its duration to that of the preceding note.

CHANGE PITCH - CHANGE PITCH randomly selects a note and changes its pitch. Again, arguments allow choosing a consonant or dissonant note and changing it to one of either kind.

EMBELLISH, REEMBELLISH, GENERATE MOTIVE - These elaborations embellish an existing motive, re-embellish a skeletal motive a different way, and generate a motive (either a new version of the original one or from a new motif), respectively. Of course, the embellishments above are used for the first two. No arguments are required for EMBELLISH REEMBELLISH; GENERATE MOTIVE has one argument to specify a new or the old motif.

CADENCE - Not truly an "elaboration," CADENCE nevertheless finds itself here because it is used to help build a phrase. CADENCE's argument is a list of the allowed cadence types for the present phrase.

SYNCOPATE - This is the first elaboration where the rhythm is changed without affecting any pitches. An argument specifies the number of tics to shift a randomly-chosen part of the passage.

PARTIAL MOTIVE - No arguments are given for this elaboration; using simple grouping and metrical rules a division in the motive is made and the first "half" is chosen.

To this list should be added "ELABORATE", because an elaborated passage can be re-elaborated (just as an embellished motive can be embellished). The process is performed simply by feeding the elaborated passage back through the elaborating mechanism. The order of these elaborations affects the result, but the program is not sophisticated enough at this time to decide on the "best" ordering (if there is one).

7.6 Conclusion

As with the music attributes (and the rhythm theory), the listing of the embellishments and elaborations given and used here is much smaller and more primitive than those that composers use, especially with elaborations. As we have defined the terms, elaboration is embellishment at a higher level. True (themetic) development is elaboration at a still higher level and with yet more sophistication in its execution. Thus while embellishment certainly has its place, it is mostly at the higher levels of thematic transformations[2] where the most creative composers do their best work. This, we hypothesize, is the right area of composition to start investigations into creativity.

[2]Embellishments and elaborations are very close to being transformations in the grammatical sense. They alter the character of an object while retaining its essential identity.

Chapter 8

AI Technology

8.1 Introduction

In this chapter we review the AI techniques used in EMOTER[1]. For each purpose or application we selected a technique we thought might be like the one a human composer would use in the same situation. When there is not enough information upon which to base such a decision, a general searching strategy tries all reasonable possibilities.

8.2 What is EMOTER?

Originally, EMOTER was intended to be a discovery system like Lenat's AM system [Davis & Lenat 82]. EMOTER would have added to its own knowledge every time it composed and would have relied on human criticism to narrow its search for great music. Instead EMOTER has become a knowledge-based system specifically for musical composition, relying heavily on internally-stored or computed knowledge-structures to help with decision-making. We do not consider it to be an expert system similar to MYCIN [Shortliffe 76] because knowledge there is embedded in rules to the extent that they are one and the same. We also do not consider it to be like Centaur [Aikins 83] because Centaur's frames and rules are too interdependent: each frame has a particular set of rules associated with it.

On the surface, EMOTER seems to share some similarities with R1 [McDermott 82]. Both are design systems that create something given a set of constraining inputs and internal knowledge. Both use generate and test. Neither use certainty factors to help make decisions. However, EMOTER separates declarative knowledge from rules whereas R1's knowledge is completely embedded in rules. EMOTER uses search strategies in addition to generate and test; R1 uses Newell's

[1] We have actually been discussing AI techniques throughout the thesis, especially in Chapter 3.

Match method [Newell 69], basically a variant on generate and test.

EMOTER does, however, have some characteristics of an expert system. Its most important knowledge structure is generated from declarative frames [Minsky 75] (i.e., the moodframes); the frames are instantiated in a knowledge structure similar to a blackboard environment [Nii 86, Erman & Lesser 75] for storing the attributes. These attributes work as constraints on the searching mechanisms described in the next section.

What, then, *is* EMOTER? It is a declarative knowledge-based system in which the knowledge is not embedded in executable rules but independently instantiated in a structure available to any operation needing it. The knowledge is kept separate from operations using the knowledge so that facts or operations can be added, removed or modified without affecting any other part of the system. EMOTER is also an interpreter of that knowledge. It takes in knowledge as data and "runs" it through a music composition "virtual machine " that outputs many correct musical solutions.

EMOTER, then, is an applications shell for music composition — neither with all its knowledge deeply embedded, nor a content-free shell with only deductive and searching mechanisms resident — but a music shell containing what we believe to be the most fundamental, universal knowledge about music (Meyer's and Cooke's theories) and the structure to apply it to almost any conventional musical style (Classical music for this thesis), given the characteristics of that style (a set of moodframes and style-specific rules) as input.

8.3 Search Techniques

8.3.1 Agenda Mechanism

An agenda mechanism [Davis & Lenat] is used to select embellishments. All legal embellishments are evaluated and ordered into a list. The best embellishment (the head of the list) is chosen and implemented. At that time some of the other embellishments associated with the note just embellished are invalidated because the new embellishment has become part of the motive and may itself be embellished. In addition, the new embellishment may have taken the only available rhythmic location between the two notes where it was inserted. Before the next embellishment can be chosen, therefore, new possible embellishments in the vicinity of the latest one implemented must be generated and inserted into the complete ordered list of possible embellishments.

8.3.2 Generate and Test

We use the generate-and-test (GAT) method only when we cannot find a more direct path to an answer. The initial choice and ordering of the motifs (described in Chapter 4) uses the GAT method, with Equation 4.1 handling the testing. If there were many possible motif combinations to work through, the method would be prohibitively computationally expensive; such is rarely the case, however.[2] GAT is also used in generating embellishments, although it is not a random generation. Every time a new embellishment is implemented, the resulting motive is tested to see how well it conforms to the desired attributes such as dissonance and angularity. Failure to conform results in more embellishment; success halts the process. In all of the uses of GAT when an ordered list of candidates is produced, the "losers" in the list are retained in case of backup.

8.3.3 Heuristic Search

The present method used to determine the best rhythm for a motif uses a heuristic search mechanism. In it, the most important pitch in the motif is paired with the most important location in the alloted space, then this process is performed recursively on the left and right sections of motif and the alloted space. The method does not guarantee finding the best answer, but it is very efficient in finding a good answer in a potentially huge search space.

8.4 AI Language Considerations

8.4.1 Why PROLOG?

Prolog's rule-based paradigm enables one to "talk about" relationships at a high level in a natural way. At lower levels of thinking, the pattern-matching (unification) and depth-first-search allow the writing of simpler code when "looking for" an answer when that answer cannot be directly computed. Unlike strongly-typed languages, Prolog demands no variable type declarations. A sorting rule, for example, may be used to sort many different types of objects. Finally, Prolog willingly prints out every correct answer it finds — with a minimum of encouragement — by exhaustively searching throught the technique of backtracking.

PROLOG's backtracking mechanism allows trial of many alternative solutions. Often no testing mechanism sophisticated enough to choose among them either exists or has been implemented in EMOTER. When this is true, all solutions are persued, generating many passages from one mood description.

[2]On another occasion we were forced to abandon a GAT routine because of severe combinatorial searching problems. The original method for determining the skeletal rhythm of a motive involved examination (generation) and ordering (testing) of every possible rhythm within the constraints of passage length and minimum duration before choosing the best one. While the method enabled a more sophisticated measurement for rhythmic stability involving durations as well as locations, the number of possible rhythms sometimes reached fairly astronomical proportions and required an unreasonable amount of time for generation *and* testing.

8.4.2 How Big is EMOTER?

The source code for EMOTER contains about 340 rules in roughly 7300 "lines" (not including documentation, which would double that figure). After loading the files, about ten minutes (wall clock time on a VAX 11/785) are required to compute the attribute values. About five minutes after that, EMOTER starts printing out phrases at the rate of one every 30–300 cpu-seconds. A typical mood specification will "generate" well over one hundred phrases. EMOTER runs on a 1 Klips PROLOG interpreter [Roach & Fowler 85].

8.5 Conclusion

We have seen how AI technology can be applied to an intuitive, emotionally-oriented, non-logical problem (music composition) and not simply return a solution (or even many good solutions), but return many good solutions for many classes of this problem[3]. The flexibility and generality that EMOTER demonstrates is due to the use of Artificial Intelligence paradigms instead of more conventional Algol-ithms[4].

[3] i.e., many different styles of music
[4] sic.

Chapter 9

Epilogue - Future Work and Conclusions

9.1 Introduction

In this final chapter we evaluate the successes and failures of EMOTER. The goal of composing sensitive, "musical" *melodies* has not been completely achieved by EMOTER as it has been developed so far. Therefore this chapter will also address areas of future research and development toward that goal.

9.2 Evaluation

How successful is EMOTER? The reader may partially judge for himself by examining the one hundred or so phrases in Appendix D. What follows in the rest of this section are the author's impressions and observations from working with the system.

Our first impression is that of amazement. From a single combination of a few moods EMOTER can usually generate hundreds of phrases with normal, human-sounding rhythms and melodic lines from a few simple abstract rules. Most of the implemented rules work very well. Melodic contour is natural, with few awkward leaps. The structure of the phrases is generally evident but not obvious — although there are exceptions. Embellishments are effective although we would like a larger set to choose from. Elaborations unfortunately do not all work — some of the most interesting ones need more debugging — but the ones that do are quite effective. Especially gratifying is the property most of the phrases have of seeming to move toward a goal or destiny. The pre-stored motifs of [Cooke 83] provide a good starting point for the phrases, but it is evident that there is some underlying theory behind his choices that is missing. Nearly all motifs are solidly built on the tonic triad, and motifs having similar contours also have similar mood-descriptions.

Had we realized it earlier, EMOTER would be generating its motifs from general, abstract rules instead of a small list of pre-defined ones. Nevertheless EMOTER has shown that out of these few motifs it can generate a plethora of different motives, each with its own individuality.

The generated phrases are occasionally excellent, easily capable of bearing comparison with human-composed phrases. We have seen no *great* music as yet from EMOTER, however, and we do not expect to because the best music attains greatness over a longer time-span than a single phrase. In addition, more knowledge about what makes great music is needed. The most common phrases produced by EMOTER are quite ordinary — certainly expressive, certainly music, but lacking some sophistication and depth. And, as the reader may notice, some of the phrases in Appendix D are quite poor. We attribute this to an insufficient number of constraints on the search space. For example, Meyer's saturation rule would limit excessive repetition. Poorly-chosen values for some of the moodframe entries are undoubtedly another cause. Compound (triple) meters such as $\frac{6}{8}$ are not rhythmically represented in a way consistent with the practice of the Classical music composers. These problems have clear solutions.

Other problems do not give way to solution so easily. Some mood combinations seem by their interaction to be destined to produce very poor phrases. Some simplifying assumptions we made simply do not work. The averaging technique, for example, often results in a value that is not characteristic of *any* of the moods which generated it. Only about half of the moods have an associated tempo moodframe, so the ones that do have a tempo often "conspire against" the others with no vote to choose a single appropriate tempo. What is needed to solve these problems is some context sensitivity among the moods in combination.

On balance, however, the phrases display surprising musicality and variety. Where does this quality come from? It is the result of EMOTER's possession of a considerable knowledge base of what makes "good" music, combined with the ability to express that knowledge through rules, tests and operations that work intimately with that knowledge.

9.3 Future Work

As we stated at the beginning, composition of truly expressive music is a difficult goal to achieve. Below we discuss some of the many possible areas of development or improvement on what we have done in this thesis. The topics range from relatively low level "touch-ups" to major new areas of exploration; we will list them approximately in that order.

1. The most obvious touch-ups are the making of many small adjustments in the attribute values, decision parameters and choices of representation methods.

2. More rhythmic/metric knowledge of compound and triple meters ($\frac{3}{4}$, $\frac{6}{8}$, $\frac{9}{8}$) is needed in EMOTER to better mimic the "dotted" rhythm commonly used with these meters. As the program stands now, complicated triplet rhythms are generated far more often than found in Classical music.

3. Relocation of a motive to another pitch-area should not be limited to a simple transposition of all its notes by the same diatonic interval. This operation forces a certain implied harmony upon the relocated motive, a harmony that might not be the one desired. In addition, there should be an operation to change the harmony of a motive without significantly changing its pitch area (e.g., C-E-G to C-E-A to change from tonic harmony to submediant harmony).

4. A more general way to describe embellishment: accented or unaccented, consonant or dissonant, pre- or post- embellished note, and single or multiple embellishing note(s). This description could be a more theoretically-sound method than the simple listing of embellishments in common practice that we use at present.

5. Composition of full melodies. When this begins to be tried, several additional ideas (Meyer's and others) will be necessary to help organize and direct them. These include the production of ambiguous passages, recognition of the role of saturation, initiation and development of musical processes such as sequences, chromaticism and more sophisticated chord progression (including cadences).

6. More sensitive evaluation of the phrase in progress would improve the method of determination of the best embellishments and elaborations to implement. The best composers are aware of everything that is happening in the music and how it affects and relates to everything else. EMOTER is crude by comparison in its knowledge of the interrelationships among the events. Sensitivity can be increased in two ways.

   (a) The music can be constantly monitored to measure the discrepancies between the desired attributes and the measured attributes of the implemented music on a global level. The discrepancies would help decide what to do next. The method we have in mind is somewhat reminiscent of GPS.

   (b) As EMOTER produces more and more sophisticated music a more sophisticated representation will be needed to store and manipulate the progress and process of composition and development. On a high level, scripts might be used to represent some of the common melodic forms. On the lower levels (motivic development and the like) a network will certainly be needed, since an interesting or creative melody follows no set structure independent of itself. As a starting point only we propose a simple representation called CONS (Constructor Of Note Sequences). It is a tree structure that includes for each of the different levels (phrase, motive, etc.) the source material, important operations performed, relationships to other parts of the structure and the resultant notes produced. A passage of music is represented in CONS as a list of items in the form (passage <level> <children> <function>), where <level> is of the form (level <phrase number> <motive number> <figure number>). If the level is a phrase the other two numbers are zero; similarly for a motive the figure number is zero. The last motive in a phrase is assumed to be a cadence. <Children> is a list of the passages and passing-tone groups in the level. The passages of a phrase or motive are <levels>. The passages of a figure are a list of pre-embellishments (if any), a pitch from the original motif and a list of post-embellishments (again, if any), all expressed in SICCLLL. A group of passing-tones is a list of the form (passingtones <passing-tones>), where <passing-tones> is a SICCLLL list of notes. <function> is of the form (<formal label> <structural function> <meaning>). The <formal label> gives a unique name within the level above. For a phrase or motive it consists of a list of a letter and a digit. For a phrase, the formal label is a capital A or B; for a motive it is a lower-case a, b, c or d. The letter refers to the origin of the item. The digit tells which version it is. The formal label for a figure lists instead of a letter and number the starting and ending locations of the notes in the figure. The <structural function> is a list describing the operations performed on the item (in the order performed). Each operation is similar in form to the representation in code of that operation. <Meaning> is represented by a word describing the function of this passage as it relates to Meyer's emotion grammar. The possible values are *establish, confirm, continue, frustrate, obscure* and *fulfill. Continue* indicates that the passage is part of a process (e.g., a sequence); the others were discussed in Chapter 2.

In EMOTER a first step toward this kind of representation has been taken, in the form of the ROADMAP facts asserted during the composing process and in two rules (PHRASE-STRUCTURE1 and PHRASE-STRUCTURE2) that determine the general phrase structure.

7. A model of human emotional response to, and intellectual organization of music would seem to be an interesting avenue of research.

8. During the course of this work, we have come to believe that there are some relationships between the theories of Cooke and Meyer. As a result, a general emotional theory of music (and other areas of art as well) could well emerge

as an outgrowth of this research.

9. The composition of music that will in every way pass for human-composed is the ultimate goal of this work. Passing this sort of "Turing test" will remain in our opinion a very difficult task indeed.

9.4 Conclusion

After listening to many of the phrases produced by EMOTER we believe that we are on the right track. Fusion of Meyer's and Cooke's ideas (within the framework of Meyer's emotion theory) into a unified emotion-mood theory is taking shape. More general, abstract rules are surfacing from the conventional music theory context-specific rules of thumb. Evidently a true theory of emotional response can be constructed.

The phrases, while not polished, nevertheless do express the desired moods, progress toward a goal, have organization, and demonstrate Meyer's emotion grammar. We have shown that good, expressive music can indeed be generated by computer using Artificial Intelligence technology. Whether automatic composition of *great* music is just around the corner or still many years away is unknown. With the experience gained building the tools and designing the techniques implemented in EMOTER we have learned enough to continue adding more musical knowledge to our musical expert system, with no end in sight. If knowledge is power then surely someday a computer program will be able to consistently generate musical masterpieces.

References

Aikins, J. S. (1983). Prototypical Knowledge for Expert Systems. Artificial Intelligence Journal, 20,2, 163-210.

Anon. (1960). Newsweek review of Illiac Suite. Jan 4, 1960, 53.

Anon. (1962). The Machine Closes In. Time, Feb 16,1962, 65.

Baroni, M. & Jacoboni, C. (1983). Computer Generation of Melodies: Further Proposals. Computers and the Humanities, 17,1-18.

Beauchamp, J.W. & Von Foerster, H. (Eds.) (1969). Music by Computers. New York: John Wiley and Sons, Inc.

Brooks, F. P., Jr., Hopkins, A. L., Jr., Neumann, P. G. & Wright, W. V. (1957). An Experiment in Musical Composition. IRE Trans. on Electronic Computers, EC-6,175.

Colby, K. M. (1975). Artificial Paranoia. New York: Pergamon Press.

Cooke, D. (1983). The Language of Music. Oxford: Oxford University Press.

Cooper, G. W. & Meyer, L. B. (1960). The Rhythmic Structure of Music. Chicago: The University of Chicago Press.

Copland, A. (1957). What to Listen for in Music. New York: The New American Library. 24-25.

Davis, R. & Lenat, D. B. (1982). Knowledge-Based Systems in Artificial Intelligence. New York: McGraw-Hill.

Ericson, R. (1955). The Structure of Music. New York: The Noonday Press.

Erman, L. D. & Lesser, V. R. (1975). A Multi-Level Organisation for Problem Solving Using Many, Diverse, Cooperating Sourdces of Knowledge. Proceedings of IJCAI-75, International Joint Conference on Artificial Intelligence, 483-490.

Faught, W. S. (1977). Motivation and intensionality in a computer simulation model. Technical report No. STAN–CS–77–633, Department of Computer Science, Stanford, CA.

Goetschius, P. (1902). Counterpoint Applied. New York: G. Schirmer, Inc.

Green, D. M. (1965). Form in Tonal Music. New York: Holt, Rinehart and Winston, Inc.

Hevner, K. (1935). Expression in Music. Psychological Review, 47, 186-204.

Hiller, L. A., Jr. & Isaacson, L. M. (1979). Experimental Music. Westport, Connecticut: Greenwood Press.

Hindemith, P. (1942). The Craft of Musical Composition. London: Schott and Co., Ltd.

Holtzman, S. R. (1981). Using Generative Grammars for Music Composition. Compu- ter Music Journal, 5(1),51-64.

Howe, H. S. Jr. (1975). Composing by Computer. Computers and the Humanities, 9(6),281-290.

Jones, G. T. (1963). Music Composition. Evanston, Illinois: Summy-Birchard Company.

Lerdahl, F. & Jackendoff, K. (1983). A generative Theory of Tonal Music. Cam-bridge, Massachusetts: The MIT Press.

Lincoln, H. B. (Ed.) (1970). The Computer and Music. Ithaca: Cornell University Press.

Lindblom, B. & Sundberg, J. (1970). Towards a Generative Theory of Melody. Swedish Journal of Musicology, 52,77-88.

Lundin, R. W. (1953). An Objective Psychology of Music. New York: The Ronald Press Company.

McDermott, J. (1982). R1: A rule-based configurer of computer systems. Artificial Intelligence Journal, 19,1,39-88.

Meyer, L.B. (1956). Emotion and Meaning in Music. Chicago: The University of

Chicago Press.

Minsky, M. (1975). A Framework for Representing Knowledge. The Psychology of Computer Vision. New York: McGraw-Hill.

Minsky, M. (1981). Music, Mind and Meaning. Computer Music Journal, 5(3),28-44.

Moorer, J. A. (1972). Music and Computer Composition. Communications of the ACM, 15(2),104-113.

Newell, A. (1969). Heuristic programming: ill-structured problems. In Progress in Operations Research, 361-414. J. S. Aronofsky (ed.). John Wiley and Sons.

Nii, H. P. (1986). Blackboard Systems, Parts I and II. AI Magazine, v. 7-2, 7-3.

Nilsson, N. J. (1980). Principles of Artificial Intelligence. Tioga Publishing Co.

Palmer, K. (1972). Compose Music (2nd Ed.). London: The English Universities Press, Ltd.

Pinkerton, R. C. (1956). Information Theory and Melody. Scientific American, 194,77.

Piston, W. (1962). Harmony (3rd Ed.). New York: W. W. Norton and Company, Inc.

Rader, G. M. (1974). A Method for Composing Simple Traditional Music by Computer. Communications of the ACM, 17(11),631-638.

Rameau, J. P. (1971). Treatise on Harmony. New York: Dover Publications, Inc.

Roach, J. & Fowler, G. (1985). Virginia Tech PROLOG/Lisp, a Dual Interpreter Implementation. Proceedings of the 18th Hawaii Interpretation on System Sciences, 2-4 January 1985, 88-92.

Roads, C. (1978). Composing Grammars. Proceedings of the 1977 International Computer Music Conference. San Francisco: Computer Music Association.

Roads, C. (1979). Grammars as Representations for Music. Computer Music Journal, 3(1),48-55.

Roads, C. (1980). Artificial Intelligence and Music. Computer Music Journal, 4(2),13-25.

Dutch, R. A., Ed. (1964). Roget's Thesaurus of English Words and Phrases. New York: St. Martin's Press.

Salzer, F. (1962). Structural Hearing. New York: Dover Publications, Inc.

Schenker, H. (1956). Free Composition. Vienna: Universal Edition.

Schoenberg, A. (1967). Fundamentals of Music Composition. London: Faber and Faber Limited.

Sherer, K. R. & Oshinsky, J. S. (1977). Cue Utilization in Emotion Attribution from Auditory Stimuli. Motivation and Emotion, 1(1977), 331-346.

Shortliffe, E. H. (1976). Computer-Based Medical Consultations: MYCIN. New York: Elsevier Scientific Publishing Company.

Toch, E. (1977). The Shaping Forces in Music. New York: Dover Publications, Inc.

Tovey, D. F. (1966). The Forms of Music. Cleveland: The World Publishing Company.

Wedin, L. (1972). A Multidimensional Study of Perceptual-Emotional Qualities in Music. Scandinavian Journal of Music, 13,241-257.

Yeston, M. (1976). The Stratification of Musical Rhythm. New Haven: Yale University Press.

Zaripov, R. Kh. (1960). An Algorithmic Description of the Music Composing Process. Automation Express,3(Nov. 1960),17.

Appendix A

A Glossary of Terms

A Word of Explanation
The terms in this glossary are defined as they are used in the thesis. In some cases a term is used in a more restricted sense than usual; in others it may have a special meaning not found outside this thesis. Occasionally we delve deeper into a subject than simply supplying a definition. Our reason is to offer more depth for the interested reader (and occasionally offer our unsolicited opinions, since they seem to be least inappropriate here).

ACTIVE TONE A pitch that has a definite tendency to be immediately followed by a more stable pitch. A member of the tonic triad is stable; any other scale degree is active and has a tendency to "move" to the closest stable pitch.

AESTHETICS Beauty. In Classical music, an "artful" balance of order and surprise, of consonance and dissonance, with the implicit assurance that all conflict will eventually be resolved.

ALGORITHMIC COMPOSITION A method of automatic composition in which the plan of composition is stated explicitly as an algorithm. With a highly structured composition this method allows the composer to "express" his composition compactly.

AMBIGUITY A lack of of coherence, ideas or a singular goal in a passage. Ambiguity creates in the listener the desire for a return to well-defined structures.

CLASSICAL MUSIC In this thesis, music of the common practice period, especially the early part of the period.

CLIMAX The goal of a passage, according to [Meyer 56]. The climax generally occurs at or near the highest pitch and the greatest dissonance.

CLOSURE Motion toward an evident goal as evidenced, for example, by one or more of the following:

- motion toward the tonic, melodic or harmonic
- rhythmic slowing down
- melodic descent
- slowing down the rate of musical process, sense of tendency, progress or development

COMMON PRACTICE PERIOD The period of western "classical" music upon which conventional music theory is based. Composers of the common practice period include Mozart, Haydn, Beethoven and Schubert.

COMPLETENESS The end result of closure:

- a return to the original (beginning) harmony
- the completion of a pattern previously made incomplete
- a pattern resolved
- a "goal" reached

CONFIRMATION The second step in meyer's emotional sequence. Established traits are confirmed by reinforcement. For example, pattern traits are confirmed by exact repetition or assignment of (some of) them to a new sequence of notes - a "new" pattern which is similar to, yet delineated somehow from the original pattern. Confirmation may occur after a break from the pattern.

CONSONANCE As used in this thesis, membership of a pitch in the current harmony. In a broader, sense it can mean an instance of adherence to the rules in force.

CONTOUR The general arrangement of pitches, harmonies, or rhythms in a passage. Melodic contour can be rising, falling, arched, wavy or level. Rhythmic contour generally increases ("more notes") as a phrase progresses. Harmonic contour can be described in terms of chromaticism and harmonic "distance" from the tonic chord.

DEVIATION A change of or from the expected or norm. A new pitch in a section is seen as a deviation (from the established pitch-set up to that time). Since deviations are most effective when or where a pattern is most complete, it follows that a climax - an accumulation of deviations - will occur fairly late in, but not at the end of, a passage. *Any* (well-formed) melody is a deviation from the "norm" of a "tonic melody" - an extended tonic chord (with no implied rhythmic content) with the tonic pitch most prominent. Rhythm (and tempo and meter) articulate - and result from - the melodic deviations, over time, from the norm, but the rhythmic element itself almost immediately becomes a norm ripe for deviation. A deviation calls attention to itself by its nature of being different from or standing out from its norm-al surroundings. The deviation encourages interest in "what's going on" - the new process expected as a result of the deviation. Establishment of the norm is a precondition for the existence of deviation by definition, and a return to the norm (which may now include norms that have been established within the melody) is necessary for a feeling of satisfaction.

DISSONANCE Non-membership of a pitch in the current harmony. The non-harmonic tones include passing tones, neighboring tones, the appogiatura, suspension, and the ornamental resolution of a suspension (echapee, cambiata and auxiliary)

DURATION An interval of time. For this thesis, an integral or discrete interval. In SICCLLL the unit of duration (and location) is the *tic*. It is the smallest allowed duration for the composition, the quantum of duration.

EMBELLISHMENT The addition of one or more structurally less essential notes to a passage. For more information see the appendix on development.

EMOTION According to [Meyer 56, p. 14], a mental state produced when a tendency to respond to a stimulus is inhibited. We believe that emotion is also generated when that tendency is finally fulfilled, and possibly when it is perceived that the goal is in sight, at the point of climax or process reversal.

"WRONG" NOTES An *occasional* "wrong" (or missing) note creates emotion, a strong yearning or expectation to "justify" it: resolve it somehow or integrate it into the music (e.g., repeat it to show it was meant to be there). This is a low-level example of the effect of breaking a rule: it creates emotion, but it must be justified.

ESTABLISHMENT The first step in Meyer's emotional "grammar". To generate expectations it is first necessary to generate something which leads one to expect a consequent. Only something familiar can lead to specific expectations (which are based on past experience), so that familiar something must be *established* before emotions can be elicited. In Classical music tonality, meter, tempo, motive and other patterns are expected to be established quickly after - if not immediately at - the beginning of the melody. A "clear" melody will establish its patterns before (and more strongly than) an "ambiguous" one. Before a pattern can be changed, it must be established - *all* aspects of the melody that are to be used as/in patterns must be established first so that their deviation is meaningful. This includes key, meter, section length and form.

FINALITY See CLOSURE.

FRUSTRATION Postponement or other inhibition of (an expected stimulus causing) a tendency to respond. Pattern traits, for example, are frustrated by changing or eliminating (or introducing new) traits of the pattern. The strength of frustration depends on the number of traits broken and the "strength" of each trait (this includes how long it has been in effect and how "obvious" it is).

GENERATE AND TEST A search technique wherein candidate answers to a problem are generated (usually in an orderly fashion) until one is found which passes some group of requirements.

GOAL (also see completeness and closure) We differentiate two types:

1. a *musical* goal can be
   - a point of relative repose - resolution of tension, low tension or the completion of a pattern for example
   - the end of a phrase, fragment or melody - a caesura near such a place creates the effect of a goal almost by default, whether or not the above is satisfied
   - fulfillment in Meyer's emotional grammar.
2. an *emotional* goal (according to Meyer) is a climax, a point of relative maximum tension.

One implies the other, both need each other, but the musical goal must follow the emotional one.

HARMONY The actual or (in this case) implied simultaneous occurrence of three or more pitch-classes (scale-degrees), or a sequence of harmonies. The scale-degrees, if sounded or referred to simultaneously, form a chord. In this thesis the chords are all strictly triads.

INTERESTINGNESS The potential ability of a stimulus to hold and keep the willing attention of an observer. An interesting piece of music has a "plan" or organization which proceeds in a fashion that keeps the listener's interest (intellectual and emotional), doesn't bore him, and doesn't lose him in excessive intricacies. The organization "makes sense" (in hindsight) in that later patterns result somehow from earlier ones, but in foresight can't be reliably predicted ahead of time. The more patterns there are at more hierarchical levels, the more interesting the piece is likely to be.

INTERVAL The vertical or pitch distance between two notes. An interval has two parts - a quality (e.g., Perfect, Major or minor) and a diatonic interval (unison, 2nd, 3rd, 4th, etc.). Unisons, 4ths, 5ths and octaves are generally Perfect (except for the diminished 5th/Augmented 4th which is also called the tritone); the other intervals are Major or minor.

MEANING In this context, the relationship of musical stimulus (sound) to the listener's reaction to it. Naturally, examples of this kind of meaning are difficult to put into words since music is a wordless medium. Other factors such as the environment of the listening experience and the listener's own previous experiences also affect the meaning of a piece of music. We have therefore simplified meaning's real meaning so as to make it amenable to investigation.

METER Ostensibly a description of the amount of duration in a measure. Meter is implicitly a description of the hierarchies of location importances within the measure. For example, the "4" in $\frac{2}{4}$ time indicates that a quarter-note recieves one beat, and the "2" indicates that there are two quarter-notes per measure. While syntactically useful, this information conveys scant semantic information. The implicit meaning of $\frac{2}{4}$ time is that there are two beats in a measure, *with the first beat receiving more emphasis than the second, and fractional beat-locations receiving emphasis directly related to their fractional value.*

METRICAL CONTOUR The metrical "shape" of a measure, i.e., the metrical strengths of locations within it where a note is allowed to exist. The first location, the so-called downbeat of the measure is always the strongest. The third beat (if there is one) is second in strength, followed by the other beats. Fractional beats are weaker still.

MOOD A feeling such as happiness or excitement. It is permanent and stable compared with emotion, which is temporary and evanescent.

MOTIF 2 to 4 scale degrees thought of as relative pitches only, with no absolute pitch or rhythm (but posessing "majorness" or "minorness" of mode). [Cooke 83] defines the ones used in this thesis.

MOTIVE A motif that has been "vitalized" [Cooke 83] with a particular rhythm and set of absolute pitches.

NORMALIZED SCALE An internal scale used for computational convenience. To avoid having to deal with different keys, a normalized major and minor scale are used exclusively. The major scale is equivalent to C major with the tonic chosen as middle C (chromatic pitch of 46). The minor scale is equivalent to A minor with the tonic chosen as the A below middle C (chromatic pitch of 43). Occassionally a "base" scale consisting of chromatic pitches zero through eleven is used. The base pitches are simply the normalized pitches mod 12.

PASSAGE A contiguous group of notes, a subset of a piece of music. A passage can be a motive, several motives, a phrase, or whatever; the only restriction on its usage is that a passage can in some way be taken as a unit.

PATTERN A passage having certain properties which are capable of being imitated. A random sequence of notes is not a pattern because it has no definable, delineable traits (except randomness).

PHRASE The smallest unit of music that expresses a "complete thought". A musical phrase is roughly analogous to a phrase in english; using the same analogy, a motive would be a word and a melody a complete sentence. The typical phrase is about five seconds long, plus or minus three seconds. It has a climax and harmonic motion ending with a cadence. A phrase is also the smallest unit of music upon which Meyer's emotional grammar can function (unless one counts a scale run with a discontinuity), but it can also function as a part of a larger emotional sequence. The number of motives in a phrase is roughly the same as the number of notes in a motive - two to four.

PHRASE-MEMBER A motif, transitional motif, skeletal motive or motive. A unitary part of a phrase.

PITCH A sound having a single predominant frequency. Chromatic pitch - a member of the complete set of chromatic pitches. Diatonic pitch - a member of the complete set of pitches in a particular key.

PITCH-PLACEMENT The average pitch and implied harmony of a passage.

PROCESS A passage which is well-defined, well-organized and predictable in that its immediate hypothetical destiny is fairly clear. Exact repetition and sequence are two common examples of a process.

RHYTHM The order of durations in a passage.

- Duple rhythm - duration ratios based on powers of two (e.g., 2/4, 4/4).
- Triple rhythm - duration ratios based on powers of three (e.g., 9/8).

Both types can exist simultaneously at different levels (e.g., 3/4, 6/8).

- Dotted rhythm - a regularly-occurring pattern of a long – short duration-pair whose total duration equals a simple metric unit (such as a beat).

SATURATION Excessive predictability in a musical passage occurring due to too much repetition or too many iterations of a process. In Classical music about three was the maximum number of iterations that could occur before the composer would depart from the pattern. [Meyer 56] devotes much time to the subject. It is one of several of his ideas that we will be doing future work with. In particular, saturation can be used to control repetition in EMOTER and maintain interest (inhibit boredom).

SCALE A subset of the twelve chromatic pitches of an octave. A *diatonic* scale in this thesis is either C-major (CDEFGAB) or A-minor (ABCDEFG with possible raised sixth and seventh scale degrees).

SCALE DEGREE A pitch viewed in the context of a major or minor scale, without regard to the actual pitch name or octave-location. Below are the names of the scale degrees in C major (uppercase letters) and A minor (lowercase letters in parentheses); in this thesis we generally use the numbers one through seven to denote them. The names can also be used to refer to the chord names "built on" the scale degrees.

- C (a) - 1 - Tonic
- D (b) - 2 - Supertonic
- E (c) - 3 - Mediant
- F (d) - 4 - Subdominant
- G (e) - 5 - Dominant
- A (f) - 6 - Submediant
- B (g) - 7 - Subtonic or Leading-tone

SEQUENCE Exact or nearly exact repetition of a motive at a pitch-level a step or two above or below the original one. Generally sequences occur in twos, threes or fours. A simple example is ABCD BCDE CDEF.

SYNCOPATION Metric dissonance. Syncopation occurs when an emphasized note occurs on a weak metric location, or when a rhythm suggests a meter different from the actual meter. An example of the former is ⊔∧⊔⊔ ⊔∧⊔⊔ (where the first character of each group is expected to be the most prominent). An example of the latter is ∧⊔⊔∧ ⊔⊔∧⊔ ⊔∧⊔⊔, with a "three" feeling superimposed upon a "four" meter. Syncopation is strongest when a note continues through a stronger metric location than the location of its initial attack. In the sample phrase, the last note in the penultimate measure occurs on a very weak metric location but is tied to the downbeat of the last measure.

TENSION In the context of Meyer's ideas, Tension is roughly equivalent to frustration. To completely describe tension in music would require writing another thesis. The opposite of tension is release or fulfillment.

TESSITURA Average pitch or pitch area.

TRANSPOSITION Uniform diatonic pitch shift of a passage with no effect on rhythm. For example, CEG transposed up two scale-degrees results in EGB.

TRIAD A three-note chord formed from every other scale degree, as CEG (the tonic triad in C major, since it is "built on" the tonic scale degree).

Appendix B

NOTES: NOTe Encoding System

B.1 Motivation

Many computer representations for music exist (e.g., DARMS, EUTERPE), but all the ones we examined are unsuitable for EMOTER because they lacked the right combination of flexibility, clarity of notation and easy adaptability to a list-oriented programming language. With the help of Dr. John Roach we designed a very flexible, list-oriented representation for conventional (non-electronic) music called NOTES (NOTe Encoding System). NOTES is ideal for use with Lisp or PROLOG since it uses the list as its organizational method. The notation is easily readable and alterable by humans, and it can be expanded with new symbols to accomodate the user's needs. It can represent and store complex passages using whatever groupings of notes or parts seem most natural. For example, a multi-part piano score can be represented as a list of chords, a list of individual lines, or a combination of the two. NOTES is *not* intended to represent any information about the physical placement of the music on the page; it only "knows" the "logic" of music.

We originally intended to use NOTES in EMOTER, but we found we also needed an internal representation that would enable easy mathematical manipulation of the notes by the program. Also, since only melodic material was to be generated, with no consideration of explicit harmony (vertical chords), orchestration or many other factors that NOTES could handle, we decided that the sophistication, complexity and flexibility of this representation was more than we needed. However, as we intend to pursue the work we started with this thesis, the advantages of NOTES will surely cause it to be used as EMOTER evolves into a more complete music composer.

The next section lists a suggested standard for NOTES. At the very least it should give the reader a "feel" for the representation. The final section lists a small sample of music in conventional notation and NOTES.

B.2 A Semi-formal Grammar for NOTES

```
NOTATION:   <> encloses a nonterminal name
            [] encloses an optional term
            {} encloses a comment
            |  separates alternative choices
            "" encloses one or more of above symbols when used in
                 NOTES; also used for enclosing character strings
```

```
<b> ::== " "[<b>]

<number> ::== <digit>[<number>]

<digit> ::== 0|1|2|3|4|5|6|7|8|9


<composition> ::== ((composer <b> <composer's name>) (title <b> <title>)
                     <movement(s)>)

   <movement(s)> ::== ( [(movement <b> <title>)] (<initial conditions>)
                        (<the part(s)>) )

      <initial conditions> ::== initial <b> (<time sig.>) (<key sig.>)
                                 [(<tempo>)] [(<movement mods.>)]

         <time sig.> ::== timesig <b> <number>/<note type>

         <key sig.> ::== keysig <b> 0 {no sharps or flats} |
                         keysig <b> <small digit><accidental>

            <small digit> ::== 1|2|3|4|5|6|7

         <tempo> ::== tempo <b> <duration>=<number> | <word>

     <the part(s)> ::== (<a part>) [<the part(s)>]

        <a part> ::== (part <b> <part name>) [(initial <b> <part mods>)]
                      (music <b> <the music>)

           <the music> ::== (<the part(s)>) {possible sub-parts} |
                            (measure <b> [<measure modifications>] <b>
                               <a measure>) [<the music>]

              <a measure> ::== (<note>) [<a measure>] |
                               % {repeat last measure} |
```

```
                    (<note group>) [<a measure>]

          <note group> ::== (<note>) [<note group>]

          <note> ::== <duration> <b> R {rest} |
                   <duration> <b> <pitch> <b> [<mods.>]

            <duration> ::== <note type> [<dots>] [<tuplet>] |
                            / {repeat last duration}

              <note type> ::== 1|2|4|8|16|32|64|128
                         {1 = whole note, 2 = half note, etc.}

              <dots> ::== .|.. {each dot increases the value
                                of the note by one-half}

              <tuplet> ::== [t<number>]
                   {e.g., three 8t3 notes = one quarter note}

           <pitch> ::== <scale degree> [<accidental>] <octave>
                   | / {repeat last pitch or rest}

             <scale degree> ::== C|D|E|F|G|A|B

             <octave> ::== <digit> {e.g., C4 is middle C, B3
                              is the B a half-step below it}

<accidental> ::== bb|b|N|#|x {double-flat, flat, natural, sharp,
                                               double-sharp}

<mods.> ::== mods. <b> <a modifier> [<b> <mods.>]

  <a modifier> ::== _ {tie to next note (must be of same pitch)} |
                 '[<number>] {begin slur identified by number} |
                 ,[<number>]  {continue numbered slur} |
                 `[<number>]  {end numbered slur} |
                 g {force notes into treble or "g" clef} |
                 f: {force notes into bass or "f" clef} |
                 c<middle-C line> {force notes into a "c" clef} |
                 ("[" [<number>] <b> <mods.>) {the modifiers apply
                      starting here, identified with this number} |
                 ("]" [<number>] <b> (<mods.>) {the modifiers
                      identified with this number end their
                      effect here} |
                 ppp|pp|p|mp|mf|f|f|ff|fff {dynamic markings} |
```

```
^ {accent} | . {staccato} | - {legato} |
">" {diminuendo} | "<" {crescendo} |
"|:" {begin section repeat} | ":|" {end section
     repeat} | (<word list>)

<word list> ::== <word> [<word list>]

   <word> ::== rallentando|allegretto|staccato|piu moso {etc.}

<middle-C line> ::== 1|2|3|4|5 {1 is the bottom staff line}
```

The user should feel free to invent his own modifiers in addition to — or instead of — these.

B.3 A Small Example

The example below can be represented by NOTES in at least two different ways: "vertically" as a chord-progression or "horizontally" as four contrapuntal lines. Listing [1] represents the passage as chords and Listing [2] represents it as lines. Note that in Listing [1] the ordering of notes within a "chord" is not important; the order of voices in Listing [2] is also not important. The only restriction is that note attacks not occurring simultaneously must be ordered correctly. This kind of flexibility of notation enables music to be "logically" represented as it is meant to be heard.

```
(music
  (measure ( (2 C5) (2 C3) (2. G3) (2 E4) )
           ( (2 D3) (2 B4) (2 F#4) )
           (4 A3 ’)
  )
  (measure ( (4 A4) (4 A3 ‘) (2. G2) (2 F4) )                [1]
           ( (4 B4) (2. G3) )
```

```
            ( (2 C5) (2 E4) )
            (4 C2)
   )
 )


(music
  ( (part 2)
    (measure ( (2 E4) (2 F#4) ) (measure ( (2 F4) (2 E4) ) )
  )
  ( (part 1)
    (measure ( (2 C5) (2 B4) ) (measure ( (4 A4) (4 B4) (2 C4) ) )
  )                                                                      [2]
  ( (part 3)
    (measure ( (2. G3) (4 A3 ') ) (measure ( (4 A3 ') (2. G3) ) )
  )
  ( (part 4)
    (measure ( (2 C3) (2 D3) ) (measure ( (2. G3) (4 C2) ) )
  )
)
```

Appendix C

Some Typical Input Mood Specifications

Here are some typical mood specifications that a user might want to input to EMOTER. As many or as few moods as desired may be entered, but there should not be any conflicting ones (e.g., both `happy` and `sad`) because EMOTER will not produce any output in such a case.

```
(static weary rigid)
(anguished incoming final dark)
(lively happy outgoing light mild)
(passive calm continuing light relaxed)
(static accepting continuing rigid)
(static sad incoming continuing dark passionate)
(tense happy longing rigid)
(passive weary sad accepting)
(happy)
(outgoing protesting tense active)
(happy calm passive)
(calm happy light)
(lively happy light)
(longing anguished passionate outgoing active protesting)
(passive calm happy relaxed mild continuing)
(mild final happy passive)
(accepting final happy calm passive)
(anguished passionate active)
(active tense sad)
(calm sad accepting final mild)
(protesting outgoing continuing)
(lively sad outgoing continuing relaxed)
(anguished incoming final dark)
```

(happy outgoing light relaxed)
(longing dark passionate)

Appendix D

A Sampling of EMOTER's Output

The following pages contain examples of phrases EMOTER generates. At the top of each page is listed the moods used as input for the phrases below. The phrases are chosen partially randomly from the first one hundred phrases generated by the mood specification. We say "partially" because we did try to include some particularly good — and some particularly poor — examples is the lists.

Outgoing, Protesting, Tense, Active

Outgoing, Protesting, Tense, Active

Passionate, Sad, Tense

Passionate, Sad, Tense

Happy, Calm, Passive

Happy, Calm, Passive

Accepting, Final, Happy, Calm, Passive

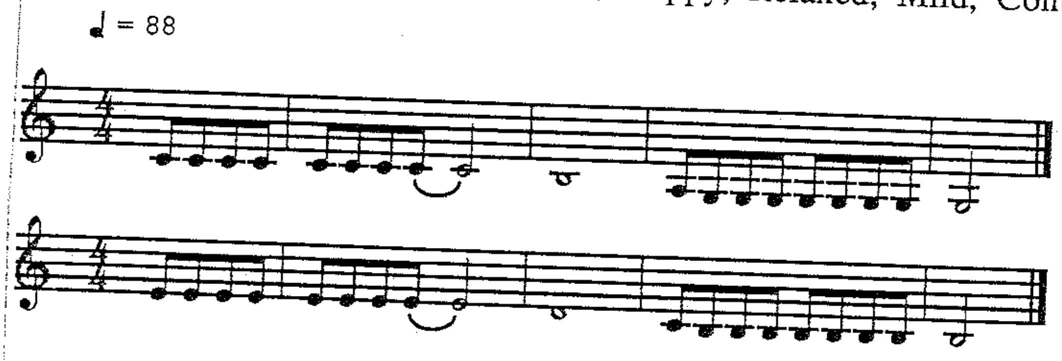
Accepting, Final, Happy, Calm, Passive

Mild, Final, Happy, Passive

Mild, Final, Happy, Passive

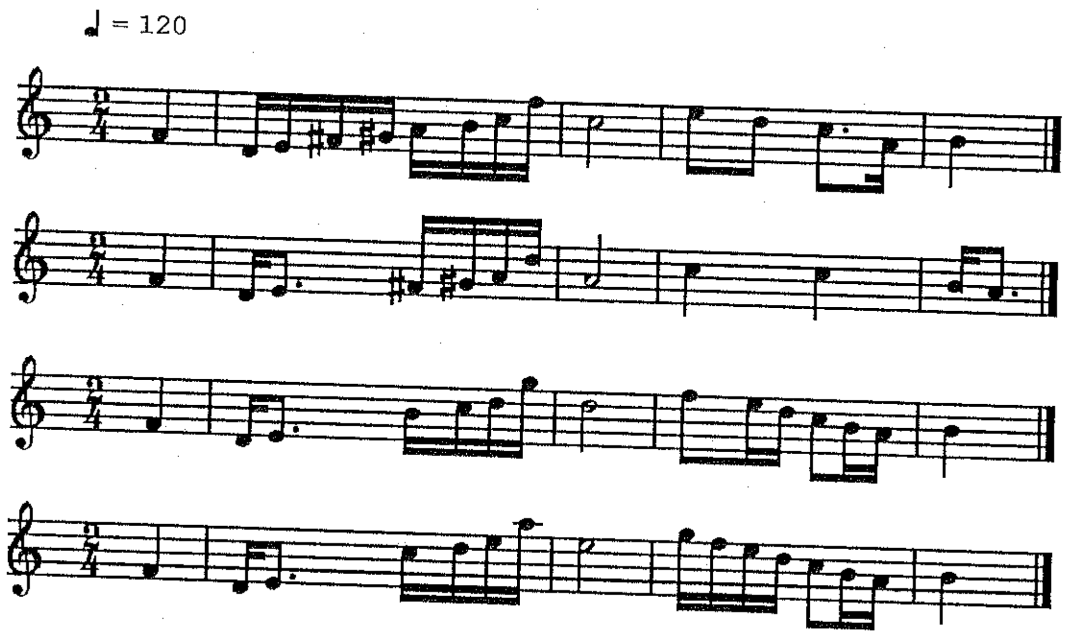
♩ = 87

Passive, Calm, Happy, Relaxed, Mild, Continuing



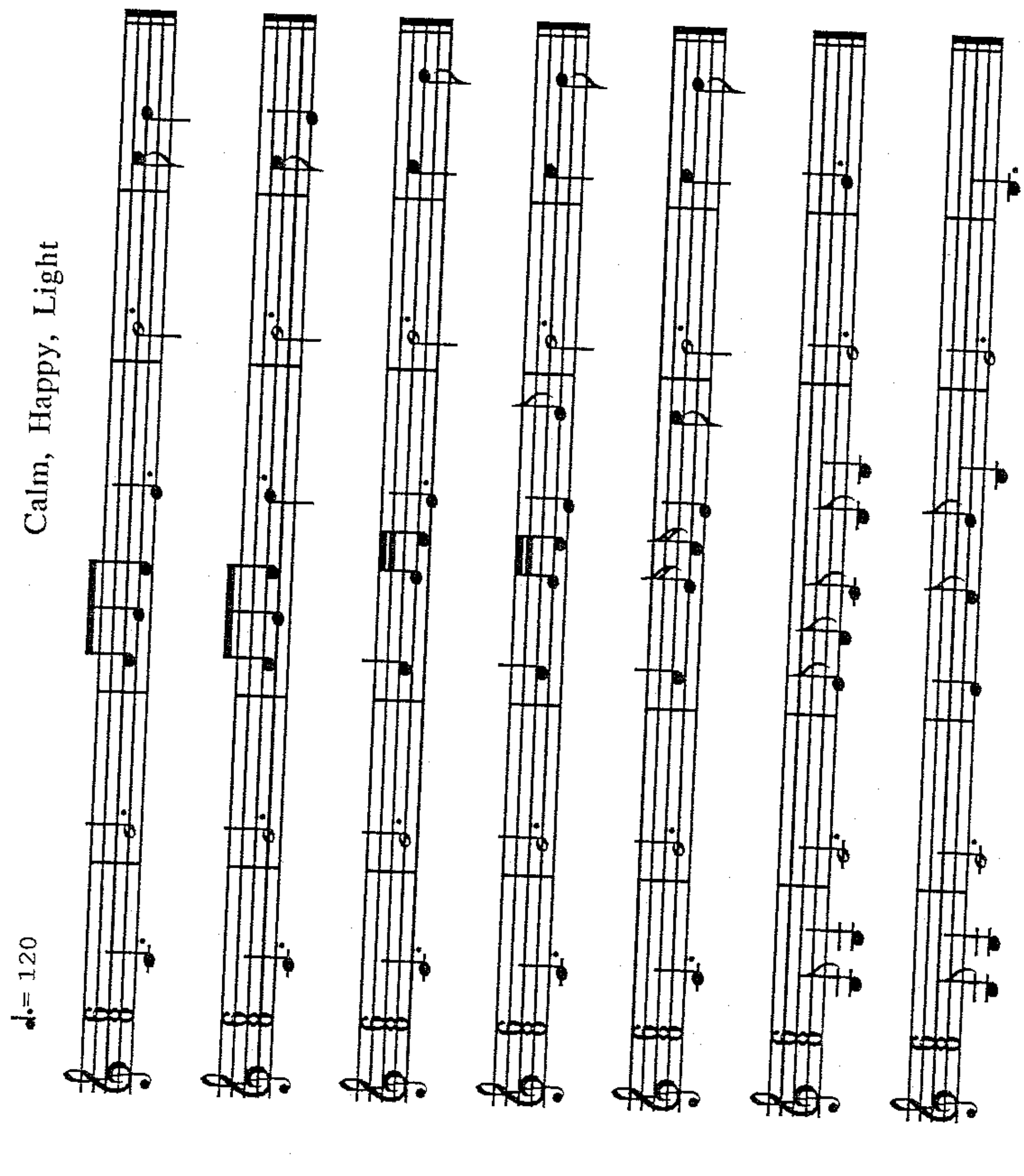
Passive, Calm, Happy, Relaxed, Mild, Continuing

Longing, Anguished, Passionate, Outgoing, Active, Protesting



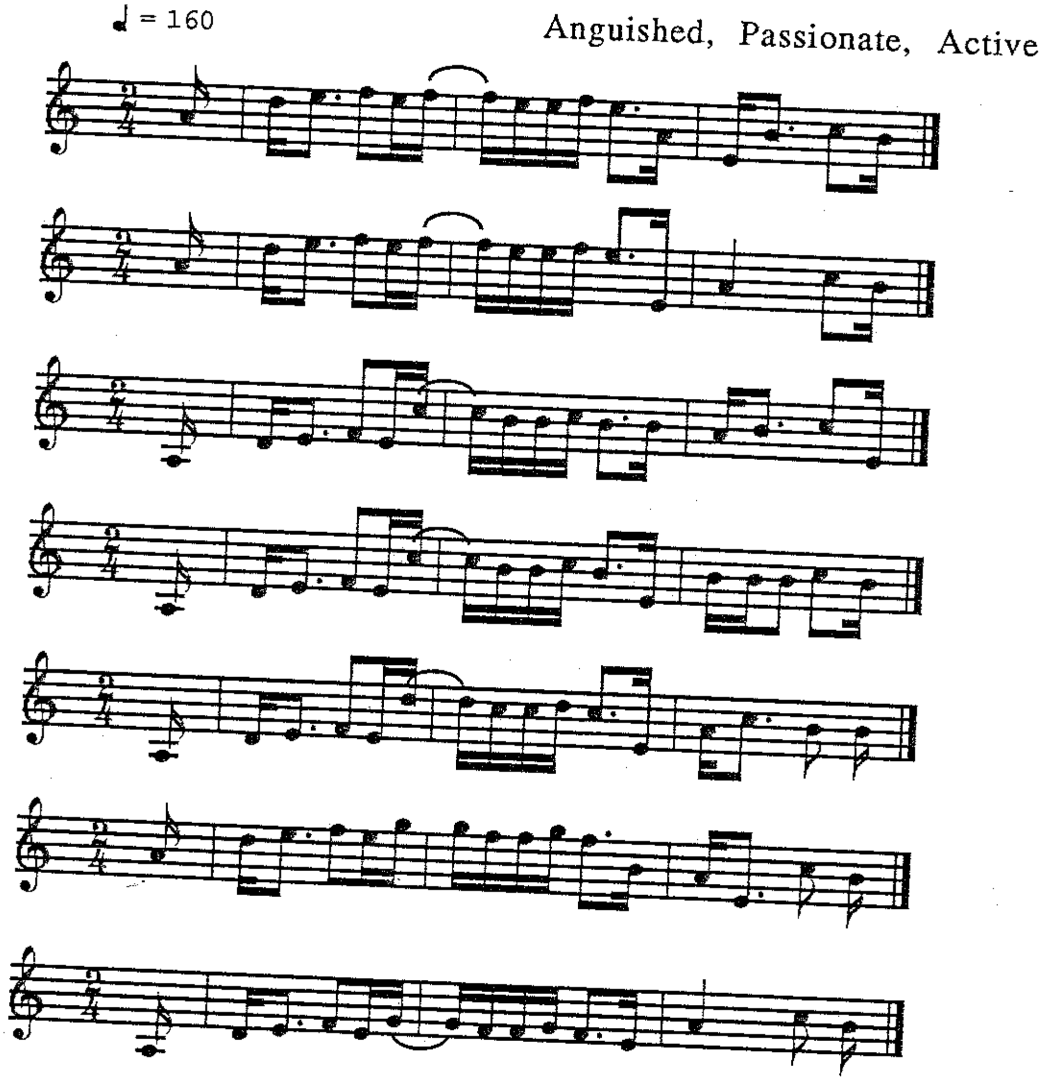
Longing, Anguished, Passionate, Outgoing, Active, Protesting

Calm, Happy, Light


Calm, Happy, Light
♩. = 120

Lively, Happy, Light

Lively, Happy, Light


♩ = 160
Anguished, Passionate, Active

Anguished, Passionate, Active