"Towards a Computer Research of Tonal Music"
Mira Belaban

ABSTRACT

Though computer assisted study of music began in the late 50's, there is yet no general, formal computer aided musical research, and existing works never claimed to initiate it.

A historical comparison of computer research on music and on Natural Languages Processing (NLP) shows that both went through a statistical-numerical period in the 50's which was followed by a "syntax first" period. In view of several basic similarities between Western Tonal Music (WTM) to natural languages one is inclined to say that perhaps, the computer study of music lacks some parallel with the current "semantics first" period of NLP, which characterizes Artificial Intelligence (AI) works on NLP. The work described here suggests a similar approach.

This paper describes the initial steps of an AI oriented study aimed at the development of a computer system for a general WTM research. The problems of level of representation and representation formalisms are treated. A discussion of possible applications and implementation is included.

Introduction

A general study of a musical style involves study of suggested theories explaining that style, their comparison, various combinations, search for new explanations, etc. The notion "musical theory" is conceived here in a very broad sense, i.e., any (partial) explanation of musical structures or styles. For example, "All melodies consist of exactly eight notes" is considered to be a musical theory as is Schenker's theory.

While the approval of a musical theory is a matter of musical evaluation and in our view, should be left to musicians, a conventional way to carry out this is through analysis or synthesis of relevant musical material, and the
presentation of the results to musicians for their evaluation. It is reasonable to expect that using the computer to perform the analysis/synthesis task would lead to a new dimension in music research, in the same way as the computer processing of natural languages opened new lines of research towards the understanding of natural languages.

The ultimate goal of this research is to develop a computer system for a general study of Western Tonal Music (WTM). This paper describes our approach and some initial steps. (For a detailed description see Malaban [1982], [1983a], [1983b], [1984].)

**General Background**

Computer assisted music research: The beginning of computer music and computer aided music research is marked by the pioneering work of Hiller and Issacson in the middle and late 50's. The early works focused on the application of either numerical or probabilistic or statistical, or information based methods to musical data base. (Hiller and Issacson [1959], Brooks et al. [1957], Pinkerton [1956]). Many works of this kind appear in Lincoln's [1970] anthology.) The main weakness of these systems was their expressive power, which was limited to that of finite automata. Hence, even simple structural properties like symmetrical segmentation of the musical material could not be handled. This period in the study of music, reminds the failure of the "mechanical translation" of natural languages, which was based on similar methods, and happened in about the same time.

A linguistic oriented period, based on elements of music theory, began about ten years later. Rothgeb's [1968] doctoral dissertation investigated the adequacy of two 18-th century comprehensive theories for the harmonization of an unfigured bass, to formal algorithmic simulation, and the applicability of some grammatical tools to theories of the tonal base. Slawson [1968] pointed at the priority of generative grammar on Markov chain models. Winograd's outstanding system in [1968] performed a harmonic analysis of tonal music pieces, using a semantic directed parsing algorithm, based on the systematic grammar model. Lídeř and Gabri's [1973] project which synthesized tonal music melodies, was based on a phrase structure grammar interrelating the notions of rhythm and pitch, and the graphic shape of a melody. Smoliar in [1976] tried to explore linguistic aspects of music programs with the aim of generalizing the results to music itself. Sundberg and Lindblom's [1976]
work also fits into this category of works, backed by computational linguistic theory. A computer study of Schenker's theory, taken as a transformational grammar, was the topic of papers by Frankel, Rosenschein and Smollar [1976], [1978], Smollar [1980], and by Kassler [1977]. In 1970 Laske conducted a research on modeling music cognition, involving ideas from the theory of generative grammars (Laske [1972]).

Most works on the computer study of music (including all of the second period works, mentioned here) focus on WTM. They share three important properties: their music scope is rather restricted, they do not claim the ability to generalize their scope, and, with the exception of Winograd's work, there is no study of the computational formalism per se. In particular, this means that there is no general, formal, computer study of WTM.

Music theory: In [1974], H.B. Lincoln wrote:

"Music theory is by no means an exact science and there is disagreement on analytical procedures, vocabulary, and varying degrees of emphasis on linear and vertical (harmonic) elements, to name but a few problem areas in this complex discipline. The capable theorist not only depends on his formal training in analysis (vocabulary, procedures, etc.) but also draws on his musical experience, intuition, and historical background - all of these being components which do not easily lend themselves to computational procedures."

(p. 286).

Indeed, there are specific, domain restricted theories that admit formal computational models (e.g. Schenker's and more recently Harms's 'Implication-Realisation' model [1977], and Lerdahl and Jackendoff's 'Generative Theory of Tonal Music' [1983]). Nevertheless, there is no generally accepted, well-defined theory of WTM.

Computer music research and natural language processing (NLP). One basic similarity of WTM to natural languages is that people that were raised on the western culture have an intuitive, unconscious (unstudied) knowledge of WTM, knowledge which enables them to accept, reject or analyze infinitely many new musical features. This similarity makes WTM especially appropriate for a scientific research since:

(1) It provides an evaluation criterion, i.e., the judgement of the "experienced listeners" (an expression taken from Lerdahl and Jackendoff [1983]), and

139
(ii) It justifies an "infinite space" requirement, i.e., a suggested model should be able to account for indefinitely large number of structures.

Considering these and many other similarities between WTM and natural languages, and in view of the intensive research on NLP, the current state of computer study of music seems rather surprising. One is inclined to say that since both NLP and the computer study of music went through a statistical-numerical period in the 50's, followed by a syntax first period, it is possible that the computer study of music lacks some parallel to the current "semantics first" period of NLP. This approach is a main characteristic of Artificial Intelligence (II) works on NLP.

Minsky [1981], Meehan [1980], Laske [1980] and Rahn [1980], all emphasize the insufficiency of syntax oriented music theories as a basis for a music understanding research, and suggest the AI approach as more appropriate.

Principles of this Research. Our basic assumption is that a system aimed for a general study of WTM requires, in the first place, a firm computational base to support representation, processing and even construction of musical theories. Observation of existing works revealed a weak computational basis in many important respects, like expressive power, simplicity, generalizability and extensibility. Consequently, we first concentrated on the study of appropriate formal mechanisms for the computer processing of WTM. It is interesting to note that a similar observation, made by Bobrow and Winograd in [1977a] led them into the design of KRL, a Knowledge Representation Language (Bobrow and Winograd, [1977a],[1977b],[1979]).

Following the general AI methodology, two main subjects of the study were:
(1) Level of representation,
(2) Representation forms.

Level of Representation

Our view is that the initial formal representation of the musical material should be as similar as possible to its statement in natural language. In that way we would achieve an initial representation which is more reliable and easy to derive. The convenience of producing a formal representation is especially important since we found that standard high level concepts are intuitively well understood but are not provided with formal definitions in text books. Observing that WTM does have a common terminology
used in statements, descriptions, and theories of WTM, it seems natural to take "the common denominator type" concepts of WTM's theories as the desired level of representation. Therefore, the common terminology found in textbooks is suggested as the vocabulary for representing WTM's theories.

The cost of the relatively easier representation for music researchers, is the need to develop a precise basis for the high level vocabulary, to enable the processing of the represented material. We found this to be a rather complicated process. It is not clear what kind of object is a musical concept; it is not clear what concepts should be considered as being of the "common denominator type"; and as mentioned above, music text books never bothered to provide precise definitions to basic concepts mainly because these concepts are very clear on the intuitive level. Thus the notions of scale, interval, musical-piece, etc. are all intuitively well understood and probably no formal definition is required in order to yield a better human understanding of them. These difficulties imply that the set of basic concepts should be studied as if it was also a musical theory. Consequently the process of establishing a high level vocabulary was split into three steps:

(a) Study of appropriate low level representation.
(b) Study of the common musical terminology in terms of the low level concepts. The result was entitled the Skeletal Theory of WTM.
(c) Embedding the Skeletal Theory in an "appropriate" representation formalism.

Low Level Representation

Our intuitive feeling is that a concept can be qualified as "low-level primitive" if it represents a "new" thing, that is, it is not a specialization of other concepts appearing in music theories and its entities are not common composite structures (like lists, sets, etc.) of entities of other known concepts. Thus, for example, we consider the concept of a musical piece as a primitive concept, but not the chord concept which is a specialization of the musical-piece concept, neither the scale concept whose entities are lists of names of pitch classes. We do not require such typical features of primitives, as independence or irreducibility. The main reason is that it seems that the concept of a musical piece, though definable in terms of notes and intervals is in some sense "most basic"; all musical structures (chord, melody, etc.) are specializations of it, and a musical theory might take it as
a basic unit, rather than as a composite structure.

In addition we expect that a considerable number of standard WIM concepts would be definable, within a single knowledge structure, in terms of the selected low level primitives, using operations over common domains like arithmetics, strings, finite lists, etc. The definition of a concept in terms of the primitives is not expected to be unique since it seems that the existence of several meaningful analyses to a musical structure or concept is a basic characteristic of music (and probably of other arts).

Based on these assumptions we suggest the concepts of Notes, Intervals and the syntactic aspect of Twelve-tone pieces as our low-level basic concepts.

Notes, Intervals: "Primitive-Elements"

The physical elementary object of music is the musical tone. The notion of distance is termed interval.

The main problems in describing these concepts in WIM are:

(a) Their multifacet appearances in musical pieces, appearances which are related to a local aspect of the piece, or to the view of the "analyst" (the emharmony feature).

(b) Ambiguous usage of each concept. For example, the word note stands for a musical tone with or without duration. Likewise, it stands for a theoretical-note-name used in tonal music terminology, again with or without duration, or for a pitch-class, i.e., a class of tones the frequencies of which differ by factors of two, or for the theoretical equivalence of pitch-class, i.e., the class of all theoretical-note-names with the same note-names, e.g., (G,♯), (♭A,bb), etc.

These problems tend to create confusion, unintended unambiguity and misunderstanding. It is unavoidable that a major portion of any work attempting to develop a formal and computer implemented study of WIM would be devoted to resolving these difficulties. We tried to achieve this by decomposing each of the concepts of Notes and Intervals into several unambiguous sub-concepts and explaining the interconnections between the new sub-concepts. This study led to the development of a unified numeral representation, entitled "Primitive Elements", for the theoretical terminology of Notes and Intervals. The Primitive Elements preserve the structure and the intended-physical-meaning of the theoretical terminology. Figure 1 describes the sub-concepts developed.
for Notes and Intervals, and their relations to the synthetic Primitive Element concept.

Figure 1.
The Primitive element concept enabled the development of a simple arith-
metics, using one additive operation, which can account for all common comput-
ations involving the WTM terminology for Notes and Intervals. In addition, 
musical structures like chords and melodies, which can be interpreted as 
either structures of notes or as structures of intervals can be described by a 
unique structure of Primitive Elements. Consequently, Primitive Elements were 
selected as Low Level Primitives, standing for both Notes and Intervals.

Twelve-tones-piece - Syntactic aspect.

In the mathematical theory of languages, the notion of a formal language 
is usually preceded by the notions of an alphabet and a string over an alpha-
bet, where a language is defined as any set of sentences over an alphabet.

In the area of natural languages, the term string is replaced by the term 
sentence. This last is generally considered as so clear and obvious, that no 
one bothers to formally define it.

Turning to the formal study of music we tend to define "the language of 
style α" as the set of all musical pieces written in style α.

Hence, the term music piece corresponds to the terms string and sentence 
of formal and natural languages. We syntactically define this term as a for-
amal object based on the theoretical terminology for notes or intervals aug-
mented with a rest symbol, accounts for the duration property of the denoted 
physical objects, and enables the distinction between successive and simul-
taneous appearances of the physical objects. In addition we want this term to 
account also for structural ambiguity.

One of the most basic phenomenon of natural languages is the ambiguity of 
their sentences, on the syntactic or the semantic level. The syntactic ambigu-
ity refers commonly to sentences which have several parsings, i.e., several 
derivations in a given grammar for the language. The ambiguity is on the 
semantic level if there exist several interpretations to an already parsed 
sentence.

Similarly, a well known characteristic of tonal music pieces is the mul-
tiplicity of their musical interpretations. The interpretations assigned to a 
musical piece, can be roughly attributed to the "musical" or "subjective" 
level of description, or also to the "structural" or "syntactical" level of 
description. The subjective level includes the whole range of personal 
interpretations revealed in the performance of a musical piece. It may be
considered as lying above the structural level since the subjective description utilizes the results of the structural description. The level of subjective ambiguity is beyond the scope of this research.

Structural ambiguity refers, informally, to the existence of several meaningful decompositions of a musical piece. It seems that achieving structural unambiguity is vital for any attempt made to formally deal with more advanced issues concerning musical pieces.

The set of twelve-tone-strings (TTS) defined below, is claimed to resolve the structural ambiguity by providing each piece with as many representations as the number of its different structural interpretations. The set is inductively constructed out of the primitive-elements set, using one basic operation, and can be generated by a context-free grammar (though not by a regular one). The representation of a musical piece as a TTS reflects, in a simple and natural manner, the intended structure.

The set of Twelve-Tones-Strings (TTS) - definition and intended meaning:

Alphabet: $I = \mathcal{R} \cup \{\phi\} \cup \mathcal{O} \cup \{\|, (, )\}$ where $\mathcal{R}$ is the set of rational numbers, $\mathcal{P}$ is the set of primitive elements and $\phi$ is a new symbol standing for a rest.

An elementary TTS is either the empty string $\varepsilon = (,\phi)$, or any pair $(p, q)$ where $p \in \mathcal{P} \cup \{\phi\}$ and $q \in \mathcal{O}$. An elementary TTS other than $\varepsilon$ stands for a theoretical name or a rest and its duration.

A TTS is defined as follows:

1. Any elementary TTS is a TTS.
2. If $\alpha, \beta$ are TTS-s and $q$ a rational number, then $(\alpha \|_{q} \beta)$ is a TTS, standing for the concatenation of the pieces represented by $\alpha$ and $\beta$ at time unit $q$ with respect to the beginning of the piece represented by $\alpha$. The ternary operation $\|$ accepts is called musical concatenation.
3. The TTS-s are only those strings which can be obtained by a finite number of applications of (1) and (2).

Example. The TTS

\[(*) \quad ((p_{1}) \|_{1} \((p_{2}/1/2) \|_{0} (p_{3}/1/2)) \|_{1/2} (p_{4})\)]

where $p_{1}, p_{2}$ are $(0, (5,0))$, $p_{2}$ is $(1, (1,0))$ and $p_{3}$ is $(0, (3,0))$ stands for the "piece":

\[\text{Diagram}

145\]
Another TTS representation for that piece is

\[
(((p_1, 1/2)) \oplus (p_2, 1/3)) \oplus (p_3, 1) \oplus (p_3, 1/2)
\]

Two useful simplifications are the horizontal and vertical concatenations, denoted by \((\alpha - \beta)\) and \((\alpha \mid \beta)\), respectively.

TTS-Representation of some typical musical structures:

1. Harmonic sentence of \(k\)-voiced choir:

\[
\bigoplus_{j=1}^{k} \left( \bigoplus_{i=1}^{j} p_{i,j} \right) \quad (n \geq 1), \quad (k \geq 1), \quad \text{where all } p_{i,j} \text{ are elementary TTS-s.}
\]

2. A melody with an accompaniment:

\[
\bigoplus_{i=1}^{p} \left( \bigoplus_{j=1}^{q} \left( \bigoplus_{k=1}^{l} q_{k,j} \right) \right) \quad (n \geq 1), \quad (k \geq 1), \quad (n \geq 1),
\]

where all \(p_{i,j}\) and \(q_{k,j}\) are elementary TTS-s.

3. A "polyphonic sentence", i.e., a sentence obtained from the simultaneous appearance of several melodies:

\[
\bigoplus_{i=1}^{k} \left( \bigoplus_{j=1}^{m} p_{i,j} \right) \quad (n \geq 1), \quad (k \geq 1), \quad (m \geq 1)
\]

The first structure describes the accompaniment as a harmonic one, while in the other, the accompaniment is formed from independent units which are basically connected to the melody.
where all $p_{ij}$ are elementary TTS-s. $k$ denotes the number of melodies, $n_i$ denotes the "length" of the $i$-th melody.

The TTS representation as input language for music: It seems that the TTS representation provides a convenient method for inputting musical information involving only the twelve tones system. Its main advantage seems to be its simple modular structure. Extra readability might be achieved by omitting parentheses from elementary TTS-s, and replacing primitive elements by their Notes' interpretation. Thus $((0,(1,0)),\frac{1}{4})$ would turn into $0 \ C \ \frac{1}{4}$, and

$$(((0,(1,0)),1)-((0,(2,0)),\frac{1}{4}))-(((0,(3,-1)),\frac{3}{4}))$$

which is much simpler. The TTS representation might be extended to account for additional information by suffixing elementary-TTS-s with additional "attributes", e.g.

![Musical notation](image)

might be represented as $(0 \ G \ \frac{1}{4} - )$, and applying the musical concatenation operation to non-notes signs like loudness signs, bar signs, verbal instructions, etc.

Example: The opening (first four bars) of Bach's twelfth choral

![Musical notation](image)

may be encoded as follows:
("Choral no. 12, Puer natus in Bethlehem" - 3) -

\[
\{(\text{2 A} \frac{1}{2} \text{I} \text{O C H} \frac{1}{2} \text{I} \text{O E H} \frac{1}{2} \text{I} \text{O A H} \frac{1}{2}) = 1\} -
\{(\text{1 A} \frac{1}{2} \text{I} \text{O C H} \frac{1}{2} \text{I} \text{O E H} \frac{1}{2} \text{I} \text{O A H} \frac{1}{2}) = 1\} -
\{(\text{1 F} \frac{1}{2} \text{I} \text{O D H} \frac{1}{2} \text{I} \text{O A H} \frac{1}{2} \text{I} \text{O A H} \frac{1}{2}) = 1\} -
\{(\text{1 G} \frac{1}{2} \text{I} \text{O D H} \frac{1}{2} \text{I} \text{O D H} \frac{1}{2} \text{I} \text{O B H} \frac{1}{2}) = 1\} -
\{(\text{1 F} \frac{1}{2} \text{I} \text{O D H} \frac{1}{2} \text{I} \text{O D H} \frac{1}{2} \text{I} \text{O B H} \frac{1}{2}) = 1\} -
\{(\text{1 E H} \frac{1}{2} \text{I} \text{O D H} \frac{1}{2} \text{I} \text{O B L} \frac{1}{2}) = 1\} -
\{(\text{1 E H} \frac{1}{2} \text{I} \text{O D H} \frac{1}{2} \text{I} \text{O B L} \frac{1}{2}) = 1\} -
\{(\text{2 C H} \frac{1}{2} \text{I} \text{O C H} \frac{1}{2} \text{I} \text{O B H} \frac{1}{2}) = 1\} -
\{(\text{1 F H} \frac{1}{2} \text{I} \text{O C H} \frac{1}{2} \text{I} \text{O B H} \frac{1}{2}) = 1\} -
\{(\text{1 F H} \frac{1}{2} \text{I} \text{O C H} \frac{1}{2} \text{I} \text{O B H} \frac{1}{2}) = 1\} -
\{(\text{1 G H} \frac{1}{2} \text{I} \text{O D H} \frac{1}{2} \text{I} \text{O B H} \frac{1}{2}) = 1\}
\]

Note that the title and the signs \( ^{3} \), \( ^{4} \), \( ^{5} \), and bar lines are also encoded.

Representation Forms

The desired nature of a formalism for representing musical theories

We expect that an appropriate formalism would possess the following features:
(a) Be powerful enough to account for significant number of interesting existing and new theories, and to support question answering about the theories.
(b) Uniform enough to enable processing by general procedures.
(c) Simple enough to ensure immediate "equivalence" of an initial formulation of a musical theory with its given verbal description.
(d) Natural enough to provide an easy transfer from natural language statement of a theory to its formal representation.

These features suggest that a representation formalism should be flexible, general, simple and human oriented; have certain amount of declarativeness and provide an explicit independent representation for musical constructs (especially those distinguished by humans). In particular, this standpoint throws away the idea of using computer programs in the initial formal stage. In fact, any procedural description, has by its nature, the flavour of a "black box" description. Another important implication is the inadequacy of known

(*) This was the main deficiency of the experiment that motivated this research where an hypothesis concerning harmonic progression was tested. (Schoenberg [1954], Balaban [1975], Sadal [1980]).
grammatical models of both formal and natural languages. For example, regular grammars can be rejected by the symmetry factor, and context-free grammars can be rejected by the reference to scale environment relation, or by the parallelism factor. However the main argument for rejecting all the grammatical models is that from the semantic point of view, the structurally unambiguous forms are usually interlaced. A similar claim was made by Narmour in (1977).

The Generalized Concept (G-C) Model

This formalism for the representation of musical theories was developed in two successive abstractions, beginning from a conceptual framework consisting of concepts, attributes and relationships, through its abstraction as a set of units called Generalized-Concepts (G-C-s), into a further abstraction as a system of constraints.

Our work on developing a formalism for representing music theories relies on the implicit assumption that it is possible to represent a significant part of music theory within a single knowledge representation formalism. This assumption reflects our intuition about the representation of musical information, and was adopted also, for its practical convenience. We don't know of any satisfactory philosophical or psychological justification for it.

1. Conceptual Framework

A good conceptual framework is one which is a natural means for describing the enterprise under consideration, which in our case is musical theories. Otherwise the very initial point of a formal research would be obscure, and the results dubious. Thus, our main guideline is to select only objects which are characteristic to human description of musical theories (and probably also to non-musical enterprises).

The suggested conceptual framework is based on the assumption that a musical theory is a set of musical hypotheses, each centered around a conceptual object. The conceptual object might be an abstract musical concept like chord basis (the set of note-names of a chord), a physical musical concept like a melody, a relation between musical objects, a particular instance of a musical concept, etc. It is our intuitive feeling that an hypothesis usually has a theme, which is the musical object it is centered around. This does not exclude, of course, hypotheses relating several objects. The object centered knowledge representation formalisms, currently very popular in AI, are all

149
based on similar assumptions. (See, for example, Bobrow and Winograd [1977a], [1977b],[1978], Brachman [1977], Minsky [1975], Schank and Abelson [1977], among many others.) Our conceptual framework is similar to Chen’s entity-relationship model [1976].

We assume only three kinds of objects: concepts, attributes and relationships.

Concepts: represent collections of musical entities which have similar structure or content. A musical entity is any distinguishable musical “thing”; it may be physical or abstract. Thus we define a concept as a set; its elements - the entities - are called instances of the concept.

Attributes: The information content of a concept is represented by a finite number of mappings from, or relations between the concept to other domains. These mappings/relations are called attributes. Each attribute is associated with one domain, i.e., a set of "values".

Concepts may be finite or infinite. In some finite cases — like the diacontic component of theoretical-note-names (see Belabas [1973a]) — the data described by the concept can be simply listed. But in most cases (even the finite), the instances and attribute-value assignments of a concept admit some regularity which can be finitely described by conditions to be satisfied. These conditions may involve other concepts and their attributes. For example, the TERTIAN-CHORD concept may be described as the set of TTS's instances of "kind vertical", which satisfy a certain “tartian structure” condition. The HARMONIC-PROGRESSION concept may be described as the set of interval names obtained between the roots of two successive tertian-chords, together with a kind attribute whose values are: (diminuendo, static, accentuated, deceptive). (See Schoenberg [1954], Sadie [1980]).

Relationships: are relations among concepts, or relations on a set of concepts. They also can have attributes, assigning values to elements of relationships.

Examples:

1) The most common relationship is the binary "is-a" relation. For example: CHORD is-a TTS, i.e., any instance of CHORD is also an instance of TTS.

2) The "consists-of" relation among concepts describes a case where the instances of one concept can be obtained from instances of other concepts, usually by means of a relatively simple function. The family of

150
consists-of relations yields a "natural-dependency" relationship on the set of concepts.

3) An important family of relationships among concepts is the "frame-of-reference" relationships, relating a concept as a possible environment or context to another concept(s). For example, the SCALE concept is a frame of reference to the NOTE-NAME concept, assigning it a degree and alteration attributes; the CHORD-BASIS concept is a frame of reference to the CHORD concept, assigning it, among others, properties of being complete/incomplete or closed/dispersed.

4) Classification relationships are 1-ary relations, intended to distinguish properties of concepts, e.g., physical versus abstract, structural versus "semantic", etc. For example, all concepts which are special cases (is-a) of the TTS concept may be marked as physical, while all the other are abstract.

II. The G-C representation model

A concept is a set of entities having similar "structure or content". This "structure or content" can be conceived as a self description of the concept, i.e., as mapping its instances into their structures or contents. This distinction between a concept's instance to its structure/content is the common distinction between symbolic objects to their meaning. Viewing a concept's "content" as just another attribute, provides extra uniformity to the concept-attributes construct since, now a concept is a set of symbolic objects, each having an associated set of attributes with one attribute distinguished as the self attribute. Moreover, all the information about a concept's instance is given by the combination of the values assigned to it by the concept's attributes (including the self attribute). Therefore we suggest to represent a concept C and its attributes by a basic-unit C, called a Generalized-Concept (G-C) which consists of all combinations of the attributes' values. That is, for a given non-empty set of attributes \( \{a_0, a_1, \ldots, a_n\} \) (\( n \geq 0 \)), where one of them, say \( a_0 \), is distinguished as the main (self) attribute standing for the "content" of the concept, and where any attribute \( a_i \) (\( 0 \leq i \leq n \)) has a domain \( D(a_i) \) of values \( D(a_i) \) is the set of "meanings" of the musical entities), the concept C and its attributes is represented by a G-C C defined as
1. \( C \subseteq D(a_0) \times D(a_1) \times \ldots \times D(a_n) \)

2. Every tuple in \( C \) describes an entity in \( C \) and a combination of values assigned to it. Thus, according to our conceptual assumptions, every tuple satisfies certain conditions which may involve other generalized concepts.

This assumption is modeled, below by the notion of a Definition of a G-C.

Turning now to relationships and their attributes, we find that they are essentially generalized concepts. However, since it is assumed that the tuples of a G-C representing a concept and its attributes are described by conditions involving (possibly) other generalized concepts, we do not see any good reason to describe a relationship with no attributes as a special G-C. On the other hand, a relationship with attributes can be represented as a G-C whose attributes are in part, attributes of the related G-C's, and in part attributes of the relationship. In the definition of this G-C, the connection between its attributes to the related concepts would be maintained by "is-a" conditions.

A conceptual description of a musical theory would probably include some redundancy, in the form of unreduced or multiple descriptions of objects.

We are unable, at the current stage of this research to decide on a "best" policy for handling redundant information. Such a policy involves decisions about computer time-space trade-offs, as well as questions about ease of translating a conceptual framework into a G-C model.

III. Definition-of-a-G-C-Theory (D-G-C-T)

This second step in the abstraction of a musical theory is based, primarily, on the assumption that concepts and their attributes can be finitey described by conditions which may involve other concepts and their attributes.

A G-C is modeled by a system of constraints called a Definition-of-a-Generalized-Concept (D-G-C). A set of D-G-C-s forms a Definition-of-a-G-C-Theory (D-G-C-T), whose "solutions", in terms of fixed points, are G-C Theories. No specific constraints language, is introduced since we believe that different theories might require different representation languages. A D-G-C is an expression of the form:

\[
C^{(1)}_{i=0} \cdots a^{(1)}_{n_k} \equiv \bigwedge_{i=0}^{n_k} A_i [C_1, \ldots, C_k] (a^{(1)}_{i=0} \cdots a^{(1)}_{n_k})
\]

152
where the $C_i$-s stand for G-C-s

the $a_j$-s stand for attributes and

the $f_{ij}$-s stand for constraints in some formal language.

The right hand side of a D-G-C is called its defining expression. A D-G-C-T is a system of D-G-C-s. An interpretation of a D-G-C-T is a many sorted, interpretation of the D-G-C-s, which leaves the Concept symbols and attribute symbols uninterpreted. An already interpreted D-G-C-T is assigned a fixed point semantics, by associating the concept symbols in the D-G-C-T with a fixed point (or fixed points) of the defining expressions of the D-G-C-T. These fixed points are the G-C theories defined by the given D-G-C-T.

The G-C formalism has similarities to several known models, including attribute grammars (Knuth [1968], Lewis, Rosenkrants and Stearns [1975]), and relational databases (Codd [1970]), and the conceptual framework reminds the Entity Relationship model for databases (Chen [1976], Ullman [1980]). However, as already mentioned the most significant similarity is to the Prototype-based formalisms, which are currently popular in AI.

A D-G-C-T for the Skeletal Theory of RTM - which is the desired (high) level of representation - was developed, using first order predicate logic language. This D-G-C-T guarantees the existence, uniqueness, and decidability of the Skeletal Theory (Balaban [1982], Balaban [1984]). However, our experience when trying to formulate even seemingly simple "theories" like the harmonic progression hypothesis (Schoenberg [1951], Sadai [1980]), or some common principles of voice leading in functional harmony, is that first order predicate logic is insufficient. Two possible extensions for future investigation are:

(a) Extension to Modal logics.

(b) Incorporating procedural knowledge.

Application: We assume that the ideas or hypotheses of a musical theory will be formulated as concepts and relationships which would be further formulated as G-C-s, and abstracted into D-G-C-s. Thus, we predict that the main uses of D-G-C-T-s would involve synthesis and analysis of instances of G-C-s through their D-G-C-s. That is, for an already interpreted D-G-C-T, the analysis process is a test if a given assignment of values to attribute symbols satisfies the D-G-C-T. (This may be generalized to partial analysis.)

The synthesis process is to find such an assignment.

153
Another way to look at the synthesis/analysis process of instances of a concept or a relationship represented by the concept symbol C with attribute symbols vector \( \mathbf{F} \) within a L-G-C-T (D(k)) is:

**synthesis:** A constructive proof of the logical implication

\[
(W ( \mathbf{F}_1 ) D(k)_{1 \leq 1}) \vdash (E \mathbf{F} C (E))
\]

**analysis:** proof of a ground case \( C' \) of \( C(\mathbf{F}) \)

\[
(W ( \mathbf{F}_1 ) D(k)_{1 \leq 1}) \vdash C'
\]

A ground case of \( C(\mathbf{F}) \) is obtained by substituting variables in \( \mathbf{F} \) by constants. These constants form the analyzed instance of \( C \).

About ten years ago Kowalski [1974] formulated, for first order logic, a constructive procedural interpretation for a resolution based proof procedure applied to Horn clauses. Moreover it was shown (van Emde and Kowalski [1976]) that for this computation model the fixed point semantics, the model-theoretic semantics and the operational semantics are all the same. In this framework, a program is a set of Horn clauses of the form

\[
A \text{ if } B_1, B_2, \ldots, B_n (n \geq 0) \text{ (or } (A \lor B_1, V \ldots V B_n))
\]

where \( A \) and all \( B_i \)-s are atomic formulas, and all free variables assume universal quantification. A computation is initiated by an initial goal statement which is of the form

\[
C_1, C_2, \ldots, C_n (n \geq 1) \text{ (more precisely } C_1 \land C_2 \land \ldots \land C_n)\]

where all \( C_i \)-s are atomic formulas and all free variables assume existential quantification. A computation is a proof resolution proof of the goal statement out of the program. The basic operation is resolution of a selected goal in the goal statement against a head of one of the clauses in the program. The constructiveness of the resolution rule provides a terminating computation with a value (binding) for the goal-statement variables, for which the logical implication program is goal statement is true. This procedural interpretation was implemented in the PROLOG programming language (Colmerauer et al. 154)
A D-G-C-T written in first order predicate logic, whose D-G-C-s are Horn clauses with head, can be considered as a PROLOG program. The analysis/synthesis process is the execution of this program. Moreover in this case, the two ways of defining the analysis/synthesis process - through fixed-point semantics and through model-theoretic semantics - turn to be the same.

Implementation: A former QLISP (Reboh and Sacerdoti [1973]) implementation of the Skeletal Theory Definition suffered from the main deficiencies of the AI languages of this kind, i.e., gross inefficiency and problems of control.

A PROLOG implementation of the Skeletal System is being currently developed. Almost half of the system has been already very efficiently implemented. This implementation supports analysis, partial analysis and synthesis of relevant musical instances.

We hope that when completed, the implementation of the Skeletal System would provide a convenient basis for the study of WTM.
References

Balaban [1975]

Balaban [1982]

Balaban [1983a]
"Towards a General Computer Study of Western Tonal Music - I", Technical Report 83-11, Department of Computer Science, SUNYA.

Balaban [1983b]
"Towards a General Computer Study of Western Tonal Music - II", Technical Report 83-12, Department of Computer Science, SUNYA.

Balaban [1984]

Bobrow and Winograd [1977a],[1977b],[1979]


Brachman [1977]

Brooks et al. [1957]

Chen [1976]

Codd [1970]

Colmerauer et al. [1972]

Frankel, Rosenstock and Smolar [1976],[1978]


Hiller and Issaassou [1969]

Yassler [1977]

Knuth [1968]

Kowalski [1974]

Laske [1972]


Laske [1980]


Lewis, Rosenkrantz and Stearns [1976]

Lidov and Gabura [1973]
Lidov, D., and Gabura, J., "A melody Writing Algorithm Using a Formal Language Model", Computer Studies In the Humanities and Social Behavior, 43(3), 135-146.


157

Meehan [1980]

Minsky [1975]

Minsky [1981]

Narmour [1977]

Pinkerton [1956]

Rahn [1980]

Reboh and Sacerdoti [1973]

Schild [1966]

Sadai [1980]

Schank and Abelson [1977]

Schoenberg [1954]

Slawson [1968]

Sollie [1976]


Smoller [1980]

Sundberg and Lindblom [1976]

Ullman [1980]

Van Emde and Kowalski [1976]

Winograd [1968]

159