Abstract Qualitative Perception Modelling and Intelligent Musical Learning

GERHARD WIDMER

Austrian Research Institute for Artificial Intelligence
and
Department of Medical Cybernetics and Artificial Intelligence
University of Vienna,
Freyung 6/2
A–1010 Vienna
Austria

gerhard@ai–vie.uucp
gerhard%ai–vie.uucp@relay.eu.net

Introduction

The research described in this article is concerned with one of the fundamental phenomena of intelligence, namely, the ability to learn. In particular, we are interested in intelligent systems that can learn, or be taught, to perform musical tasks and solve musical problems. This kind of research naturally leads to the question of what the prerequisites for effective learning are. In this article, it will be argued and demonstrated that fundamental knowledge about the domain is important and, indeed, indispensable if a system is to learn problem solving rules for a complex musical task effectively. That again leads us to ask what fundamental musical knowledge is and how it can be represented and reasoned about in a computer program. After all, the musical *a priori* knowledge we want to endow our program with should not be an ad hoc concoction of bits and pieces designed just so as to make the particular learning task feasible; that would clearly constitute a case of cheating and would certainly not advance our systematic understanding of music and musical cognition. Rather, it should be 'natural' and psychologically plausible in the sense that we can expect any human listener who has grown up in the western musical culture to possess, however tacitly, the same knowledge or intuitions. That will place this kind of research in the realm of what Otto Laske has called 'Cognitive Musicology' (Laske 1988) by making possible the experimental validation of abstract theories of musical perception and knowledge. In the context of learning to solve specific musical tasks, the adequacy or inadequacy of a general model of musical knowledge should materialize in the system's ability to easily and 'naturally' learn certain concepts and maybe have difficulties with others. Such an approach would thus enable us to demonstrate the adequacy of a general musicological model in a convincing way, and also to pinpoint weaknesses in the model that might otherwise be hard to identify.

The research described in this article is a logical continuation of an earlier project that dealt with much simpler musical problems (see Widmer 1991c). It contributes to the above–mentioned goals in several ways: the article presents a general, abstract perception model for a comparatively simple sub–domain of tonal music; the model is meant to capture, in a qualitative way, some of the aspects that govern the way people 'hear' harmonized melodies. This is the fundamental musical 'knowledge' we want to equip our system with. We then present arguments for the importance of such knowledge for learning and describe a system that uses the perception model in the process of learning to harmonize given melodies from examples of correct and incorrect solutions. Accordingly, the article is divided into two parts; the first part is devoted to the general perception model; the second part then describes the learning system and illustrates its workings with an example.

Both the qualitative perception model and the learning system have been implemented and tested with various selections of pieces. One such experiment is briefly described in the appendix. One of the contributions of this article that we would like to stress is that it demonstrates a technique for modelling perceptive phenomena at a very abstract level by using formalisms from the field of qualitative modelling. This seems to us a very promising research direction for the future, both for theoretical and practical musical applications.

Qualitative Perception Modelling

What do people know about (tonal) music? How do they 'hear' and 'understand' music? And how can a machine 'know' about and 'understand' music? The question how music 'works' has been the subject of music theory and psychoacoustics for a long time. Especially psychoacoustics would appear to be a promising basis for a model of musical hearing. However, for the purposes of a project such as ours, the level of psychoacoustic phenomena is too detailed and too remote from the phenomena we want to model; it simply cannot serve as a tractable basis for an AI system that is to learn high–level tasks like harmonization. We must find a level of abstraction that combines transparency and simplicity with a miminum (and hopefully more than that) of musical and psychological plausibility. Also, systems that are supposed to interact with human musicians or listeners must have the ability to communicate with the user in intelligible terms, which again presupposes the command of a language and terminology that is meaningful to humans.

For the purposes of this project, we have therefore adopted a notion of musical understanding similar to the way Lerdahl and Jackendoff define musical understanding in human listeners (Lerdahl and Jackendoff 1983). That is, a system is said to understand (some aspects of) a piece of music if it can organize the piece in a coherent fashion so that what was at first an unstructured sequence of pitch events (notes) becomes a meaningful structure that can be reasoned about and discussed in abstract terms. As a consequence of this, the system will dispose of notions such as *structural salience*, *tension*, *relaxation*, *smoothness of progression*, and the like, and will be able to communicate its understanding of a piece of music in such abstract terms – terms which are also familiar to human music listeners.

A Qualitative Model of the Perception of Harmonized Melodies

The model to be presented is meant to be a structured hypothesis about the perception of simple harmonized songs (i.e., melodies accompanied by chords, as in the popular song literature for guitar players or the by now classical 'Real Book' for jazz musicians, where chord symbols are associated with the notes of a melody). It will specify the main sources of musical information that can be extracted from a piece, the types of information that can be extracted, and how these various items of information combine to yield increasingly abstract 'feelings' about a given piece. In order to prevent any possible misunderstanding, it should be stressed here that we are not trying to simulate real−time perception; for the moment, we are content with modelling a kind of *a posteriori* understanding of a piece after it has been heard. Also, the perception of rhythm as a separate phenomenon is neglected; rhythm enters into the model only insofar as it plays a role in determining the relative structural importance of musical events. And finally, in the current phase of our research, we do not try to model the individual listening history of a human listener, with all its related implications.

Qualitative models

As mentioned above, our project calls for an approximate, abstract level of modelling. In Artificial Intelligence, various techniques of *Qualitative Modelling* have been developed in order to be able to model systems that are too complex (or not even exactly known) to be described and simulated in detail. Originally, qualitative models and formalisms were mainly intended to describe physical or technical systems and processes; the most prominent approaches are *Qualitative Process Theory* (Forbus 1984), *Qualitative Simulation* (Kuipers 1986) and DeKleer's *Theory of Confluences* (DeKleer and Brown 1984). Here we will show how some aspects of musical perception can be modelled in a similar way, and that this is indeed a very natural way of expressing musical intuitions.

A *qualitative model* or *theory* of some problem domain is a structured model that specifies the most important parameters/entities/concepts of the domain and the most important relationships among these, but only in a qualitative way. Qualitative models are by necessity an abstraction of the real world. They do not deal with precise quantitative values; quantities are usually described only through qualitative terms like *high*, *low*, *increasing*, *decreasing*, etc. Also, qualitative models usually specify only *dependencies* between different parameters, but not exact relationships. In addition, such models may be incomplete (missing some relevant concepts and/or relationships), incorrect (containing statements that are not true in the 'real world'), and refer to observable as well as to non−observable variables/concepts. All these features make them an elegant tool for modelling imprecise and incomplete knowledge, but also a poor basis for real problem solving. That is where the need for learning comes in. The second part of this article will have more to say on that.

In our qualitative model of music perception, the basic representational concepts will be musical events (notes, chords, or some combination of these), perceivable effects (sensations evoked in the listener by some features of a musical event) on various levels of abstraction, and influences between effects (statements that indicate how one effect can contribute to the emergence of another, more abstract, effect). The strength of these effects will be measured in qualitative terms only; for instance, the degree of contrast in some musical passage will be said to be either *extremely_high*, *high, moderate*, *low*, or *extremely_low*.

Influences between effects will be specified in an approximate way only, via qualitative proportionality relations.

Basic musical perceptions

So let us see what kinds of things we perceive when we listen to a harmonized melody, and how this may give rise to general feelings of pleasure or displeasure with a heard piece.

The first thing we note is that the structure of music is multi–dimensional, and that audible effects can arise on any of these dimensions. In the case of harmonized songs, at any point in a piece we may hear and consciously perceive some feature of an individual event, be it a single note or a single chord, or we may experience effects that are due to some combination of events; this applies to sequences of notes in the melody (*melodic dimension*), sequences of harmonies/chords (*horizontal harmonic dimension*), and to the simultaneity of notes in the melody with accompanying chords played (*vertical harmonic dimension*). In addition, we may more or less automatically attribute metrical stress to different events, we hear motivic relationships between various parts of the melody, we perceive certain events to be more important (salient) than others – for instance, almost everybody will perceive a neighboring note or a passing tone as subordinate to its surrounding events –, and finally, people have strong intuitions about the possible tonalities implied by a melodic segment (even if they possess no vocabulary for expressing these intuitions). Again, this list is incomplete; it omits factors like perception of timbre or the influence of stress and the like. Our system deals with printed music only (for obvious reasons).

While it is, and will continue to be, a matter of debate exactly how listeners manage to get some idea about things like implied tonalities or metrical and grouping structure, it seems safe to assume that they do it somehow. When defining our qualitative model, then, it matters not so much how we decide to implement some perceptive skill, but rather that the kinds of perceptions we integrate in the model are psychologically and musicologically plausible. For instance, psychological experiments indicate that people do indeed have some notion of relative distance (along the circle of fifths) between harmonies – see, for instance, some of the experiments cited in (Bharucha and Todd 1989). While it is unreasonable to expect that most people explicitly know something about the circle of fifths, it seems perfectly legitimate to realize harmonic distance perception directly via the circle of fifths in our model.

The above list of effects, however, is not the whole story. In addition to such local phenomena, we experience more abstract sensations when listening to a piece. For instance, a succession of musical events that are very different in some respect may give rise to a feeling of *contrast*. Altered notes, harmonies moving away from the tonic, or chords involving factors other than the basic triad create a sense of *growing tension.* Harmonic transitions may be felt to be *smooth* or *abrupt*, and the like. Even more abstract sensations can emerge from these effects. For instance, *contrast*, *diversity*, and the buildup of *tension* are factors that contribute to an overall feeling of *interestingness* of a passage. It seems to be, at least partly, on the basis of such abstract effects that we make intuitive judgments concerning the 'goodness' of a piece. Levels 4 and 5 of our model (see Fig.1) list some of these abstract sensations. The list is based mostly on introspection, and it is by no means complete. Suggestions for additions are welcome.

The full model

We are now in a position to present the model in its full complexity. It is shown in Fig.1. Note that this is not a process model; the various boxes and connections between them are simply meant to represent different classes of perceivable effects, corresponding to the different dimensions mentioned above, with arrows indicating the direction of influences between effects.

The basic structure is as follows: levels 0 through 6 describe increasingly more abstract views and/or perceivable effects of a harmonized piece. (There is no special meaning to the exact number of levels. Also, the abstraction hierarchy is not really strict; for instance, the *vertical harmonic dimension* view (level 3) is in no sense more abstract than, say, the *horizontal harmonic dimension* at level 2; it just happens to take into account both melody and harmony, which is why it was assigned to level 3.)

Level 0, the harmonized piece itself, is the object level. It represents the thing to be analyzed. Levels 1 and 2 are separated into groups of *melodic viewpoints*, which take into account only features of the melody, and *harmonic viewpoints*, which deal with qualities of the underlying chords only. These viewpoints are somewhat reminiscent of Ebcioglu's *views* (Ebcioglu 1988).

Level 1 depicts the most local and isolated features that can be perceived; it is the level of individual events. Looking just at the melody (*melodic viewpoint*), the individual events are the single notes, and each of these has some intrinsic qualities that can be perceived. For instance, *pitch height* is sometimes consciously heard, especially when an extremely high note is played. Altered notes can have a strong *leading tone quality*, and so on. But more than being an effect category in itself, the class of note features is an important input to the time–span reduction and tonality analysis level (level 2). For instance, the *duration* of a note plays a role in the determination of *meter* or the relative *structural importance* of the note.

On the *harmonic viewpoint* side, the individual events at level 1 are the single chords. Here again, each chord in isolation has features that might be perceived as such (e.g., the degree of *intrinsic tension*), but more importantly, these effects contribute to the emergence of more abstract effects, when events are viewed (or heard, rather) in the context of other events (see level 2).

Level 2 is the first level where sequences of events are considered. On the *melodic viewpoint* side, hearing notes in the context of other notes enables a listener to infer the *metrical structure* of the piece, to segment the stream of notes into groups (*grouping analysis*), to hear *motivic relationships* between groups, and to assess the *relative importance (salience)* of events. These perception skills (excluding motivic considerations) are realized in our model by a component that performs metrical analysis, grouping analysis, and plausible time–

span reduction as described in (Lerdahl and Jackendoff 1983). This component extracts information about the *metrical strength* and *relative structural salience* of notes in the melody, which can influence the perceived prominence of most of the more abstract effects, as indicated by the bunch of arrows emanating from the time–span reduction box. Also, at the level of melodic segments, the listener can sometimes infer *implied tonalities* and *local key shifts*. This is modelled by another component that employs some well–known heuristics for guessing tonality (see, e.g., Longuet–Higgins 1988; Sloboda 1985).

On the *harmonic viewpoint side* of level 2 (*horizontal harmonic dimension*), the events that are analyzed are sequences of chords, with perceivable effects including the *harmonic distance* (along the circle of fifths) between successive harmonies, differences in intrinsic tension between neighboring chords (which creates effects of *local tension increase* or *decrease*), and the like.

Level 3 depicts those viewpoints that combine aspects of both melody and underlying chords. The *vertical harmonic dimension* creates effects that derive from simultaneously hearing some note(s) in the melody and an accompanying chord. Here, the degree of *relative consonance* between the note(s) and the chord is particularly important.

The second dimension at level 3 is the *dimension of implied (local) key*: given that we are able to extract tonality information from the melody, we can hear and judge each chord in the framework of the hypothesized local key. For instance, we can sense a chord to be more or less *stable* within the current key. This view uses information from the *tonality analysis component* (level 2) and also knowledge about *intrinsic qualities* (level 1) of the chords under scrutiny.

Finally, levels 4 and 5 are the above mentioned abstract sensations that emerge from the combination of lower–level effects and, in turn, more or less directly influence our overall judgements of the 'goodness' of a piece (level 6). The arrows coming from below indicate which classes of effects can contribute to these abstract sensations. (The effects *diversity* and *harmonic rhythm* are of a more global nature than the others at this level; they are not dealt with in the current version of the model.)

Formalization of the model

What remains to be given is a more detailed discussion of the meaning of the arrows in Fig.1. After all, they determine what the model can be used for. The arrows describe possible influences between effects. That is, they describe in which way one effect can contribute to the emergence or prominence of another effect. Basically, there are three types of links:

Positive qualitative proportionalities (Forbus 1984):

The statement "effect E1 is positively qualitatively proportional to effect E2" – written as Q+(E1,E2) – may best be paraphrased as "if E1 increases, E2 will also, all other things being equal". Q+(E1,E2) does not mean that E1 is the only factor on which E2 depends, nor does it mean that E1 must always be present when E2 is.

An example of such a proportionality is

Q+ (consonance_note_chord, harmonic_stability)
("consonance between a note and a simultaneously played chord has a positive influence on perceived harmonic stability")

Negative qualitative proportionalities:

A statement "effect E1 is negatively qualitatively proportional to effect E2" – written as Q–(E1,E2) – means, in analogy to Q+, "if E1 increases, E2 will decrease, all other things being equal".

An example of such a proportionality is

Q– (chord_distance, smoothness_of_change)
("the greater the distance (along the circle of fifths) between two adjacent chords, the less smooth will this transition be perceived")

General qualitative dependencies:

The statement "effect E1 (partly) determines effect E2" – written as D(E1,E2) – means that there is some partial functional relationship between E1 and E2, or, informally, "same values of E1 will tend to go along with same values of E2, and if E1 changes, E2 will also change somehow, all other things being equal". This is obviously the weakest type of link, for it tells us nothing about the exact relationship between E1 and E2; all we know is that E1 is somehow relevant to E2.

An example of such a dependency is

D (contrast, interestingness)
("local contrast is one of the factors that may be related to the 'interestingness' of a passage; which forms (and degrees) of contrast make a passage more or less interesting is not known in advance")

It is exactly the vagueness of the relationships Q+, Q–, and D that makes them ideal for expressing musical intuitions: they allow us to put down our intuitions in all their generality and uncertainty; they do not force us to write down statements that pretend to be more 'precise' than our own understanding.

Taken in its entirety, the model seems to give a good account of at least some of the factors that underly our hearing of harmonized melodies. Again, the claim is not that the structure of the model mirrors the exact structure of human perception processes, let alone that ordinary listeners dispose of the kind of vocabulary used on the various levels of the model (especially the more 'technical' levels 1 through 3). What is claimed here is that listeners do somehow manage to extract information from a heard musical surface in order to arrive at 'intuitive' judgements concerning the quality of a piece, and that the information they extract must be something like the effects given in the model. And even if we were wrong in this claim, the model will still allow a computer to make intelligent judgements about simple music and to communicate its understanding of a piece to human musicians in intelligible terms.

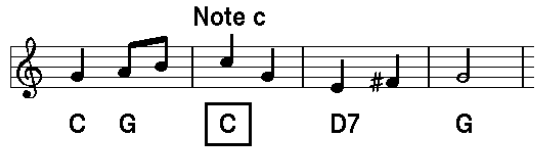
Using the model to explain musical situations

Generally, the model supports several different types of logical inferences (Widmer 1990). What is relevant in our context is its explanatory power: the model can be used to explain or verify why some qualities of a musical situation should be perceived. Take, for instance, our familiar line from Fig.1 and ask the system to verify that the C major chord at the beginning of measure 2 is indeed heard as a good harmonization. Given the goal to explain

good(C, c)
(”the chord C is a good harmonization for the note c in the melody”)

the system will find the hierarchical explanation sketched in Fig.2 by traversing the model top–down and recursively verifying the preconditions for the current explanation goal. It succeeds in showing that the C major chord sounds good because it strongly contributes to satisfying the listener’s longing for coherence of a musical passage. (The parenthesized numbers to the left of the explanation structure indicate the level in the qualitative model (see Fig.1) where the corresponding effects or concepts are defined.) Fig.2 only sketches the structure of the explanation; explanations constructed by the system are considerably more complex and carry more information than shown here. We will encounter an explanation in its full complexity later in this article.

Although this explanation may look as if it were based on sound deductive arguments (the categorical ”because”), it must be borne in mind that it is, in fact, only a plausible explanation (cf. Collins and Michalski 1989): in the qualitative model, most influences between factors at different levels are given only as qualitative proportionalities, which cannot serve as the basis for strict proofs. But given a fact to explain, the system can look at all relevant features of the situation and assess the relative plausibility of different possible qualitative explanations. In our case, the explanation given in Fig.2 is indeed very plausible, since there are no contradictory factors.




```
;;; Top 2 levels of explanation (see model in Fig.1)

(6)  good( C, c)   because                                         ;;; C is a good harmonization for note c because
(5)        coherence( C, c, extremely_high)   because              ;;;  C creates a sense of coherence because
                   previous−chord( C, G)                           ;;; chord C is preceded by chord G
(4)          and   tension−relaxation( C, c, moderate)             ;;;  it moderately relaxes some existing tension
(4)          and   smoothness−of−change( G, C, extremely_high)     ;;;  and the change from G to C is smooth
(4)          and   harmonic−stability( C, c, extremely_high)       ;;;  and the chord is highly harmonically stable


;;; Structure of complete explanation

(6)  good( C, c)   because

(5)     coherence( C, c, extremely_high)    because

             previous−chord( C, G)                                 ;;; chord C is preceded by chord G in the piece

(4)       and tension−relaxation( C, c, moderate)    because
(2)                local_tension_decrease( G, C, moderate)    because
                          global−key( c_major)                     ;;; there is some moderate relaxation of tension
(1)                   and distance_from_tonic(G,c_major,1)         ;;; because the C chord is closer to the global
(1)                   and distance_from_tonic(C,c_major,0)         ;;; tonic than its predecessor G
                      and 1 > 0
                                                                   ;;; the relaxation effect is perceived even more
(2)            and salience( c, high)    because                   ;;; strongly as the point where it occurs (note c)
                          metrical_strength(c, high)               ;;; is of high metrical and structural salience
                      and structural_salience( c, high)            ;;;  (determined from time−span reduction)

(4)       and smoothness−of−change( G, C, extremely_high)   because    ;;; the transition from G to C is
(2)                relative_chord_distance( G, C, low)    because      ;;; smooth because G and C are
                          distance_on_circle_of_fifths( G, C, 1)       ;;; just one perfect fifth apart

(4)       and harmonic−stability( C, c, extr._high)    because       ;;; C is highly harmonically stable because
(3)                relative_consonance( C, c, extr._high)    because   ;;; 1.: the note c is consonant with
                          scale_contains_note( c_major, c)             ;;;     the scale implied by the chord
                      and chord_contains_note( C, c)                   ;;;     and with the chord itself

(2)            and plausible_local_key( c, c_major)                    ;;; 2.: C is stable  within the
(3)            and stability_of_chord_in_key(C,c_major,extr.high)   because    locally implied key (c major)
(1)                       tonic−chord( C, c_major)                     ;;;     because it is the tonic chord
(1)            and intrinsic_stability_of_chord( C, high)    because   ;;; 3.: and in addition, as a
                          chord_type( C, triad)                        ;;;      root position triad, the
                                                                       ;;;     C chord is intrinsically stable
```

Fig.2: Explanation why chord C is a good harmonization for beginning of measure 2

Learning to Solve Musical Problems

We are now in a position to describe how the general musical knowledge can be put to use in a system that learns to solve specific musical problems. Let us start with a brief description of the learning problem. The system's task consists in learning rules to harmonize given melodies, i.e., associate chords with certain notes in a melody, based on solutions (harmonization examples) provided by a teacher.

When trying to harmonize a new melody, the problem solving component of the system (see below) processes a given melody from left to right, checking for each note whether it calls for a change of harmony, and if so, deciding on the best chord to place at this point. Accordingly, there are two classes of rules to be learned (examples of rules of both types can be found in the appendix):

Rules of the type harmonic_change(LastChord, CurrentNote) IF <Conditions>.
These rules state conditions under which the current note in the melody calls for a change of harmony, i.e., it should no longer be accompanied by the last chord, but a new chord/harmony should be introduced at this point.

Rules of the type good(Chord, CurrentNote) IF <Conditions>.
These rules suggest possible chords that might go well with the current note in the melody. They should take into account both melodic and harmonic considerations.

The condition parts of both of these types of rules can potentially refer to any observable surface feature of the musical situation. In addition, they may make reference to some higher−level concepts that are defined in the qualitative model. For instance, conditions on the metrical strength or structural salience of a note may also appear in a rule. The representation language for rules is thus very rich, and the number of possible rules that could be formulated in this language (and hence would potentially have to be considered by the learning system) is enormous. The following sections show how this intractable space of possibilities can be radically constrained by using the general knowledge represented by the perception model.

Knowledge−intensive learning and the perception model

Ideally, one would like to design systems that learn to solve problems and deal with new situations simply from experience, for instance, by watching some human expert solve problems in some domain. The learning strategy would be empirical generalization (induction) from specific observations (see, e.g., Michalski 1983). Although such an approach is possible in principle, the practical problems are enormous. They derive from the problems inherent to empirical induction: the number of possible generalizations that can be drawn from a finite set of observations is usually extremely large, and without any additional knowledge about the domain, a learning system has no way of choosing between these.

And we, as humans, certainly don't learn on a purely empirical basis. We know a lot about the world, and what we know (or believe to know) strongly constrains the inductive generalizations we are willing to accept as plausible. This allows us, in many cases, to focus on the most plausible hypotheses and thus learn more effectively, committing fewer errors. This holds for music as well. As we have seen, ordinary listeners possess general musical knowledge in the form of habits of perception, and this knowledge influences the way new pieces are judged. Here we will show that this knowledge can also help in the process of learning new musical rules: the learning system uses the qualitative perception model as a guide in the learning process.

It might seem that the learning task becomes trivial once this perception model is available: for instance, the system already has some notion of implied harmonies or local key. There are, however, at least two arguments to be made in refusal of this view:

Firstly, actual harmonization requires decisions that are outside the scope of the model (in fact, any kind of strategic decision is outside the scope of the model). This includes decisions such as when to switch to a new harmony (chord), how to choose among several possible local solutions, how to deal with situations that require the violation of some of the constraints specified in the model, and how to negotiate between local and global considerations to arrive at a coherent, well–balanced solution. Experience with specific solutions presented and classified by a teacher is needed to learn rules for such decisions.

Secondly, even if the knowledge contained in the perception model were complete in the sense that the 'correctness' of every harmonization could be completely verified on the basis of the model, for the system to become an effective problem solver it is necessary that the relevant knowledge be represented in the form of rules that can be directly applied to a particular problem state in order to make a decision as to how to proceed. In other words, the knowledge must be *operational*. The qualitative model does not meet this criterion. Since it is a model of perception, the knowledge could only be used in a generate–and–test manner: randomly generate a harmonization and then check whether it sounds acceptable. (Note that this corresponds closely to the general human listener situation: ordinary listeners usually have a 'passive' knowledge of music. For instance, they may hear that some harmonization sounds bad, but not necessarily be able to say what the correct harmonization should be.)

In the field of Machine Learning, the problem of learning operational rules from non–operational knowledge has been extensively studied, and the most well–known approach has come to be known as *Explanation–Based Learning* or *EBL* (Mitchell et al. 1986; DeJong and Mooney 1986). An EBL system learns rules from training examples in three steps: first, it constructs an explanation (= a proof) on the basis of its general knowledge that an example is indeed a valid instance of the concept or rule to be learned; this explanation is then generalized so that it holds for a whole class of situations; and finally, the generalized explanation is compiled into an operational rule that applies to exactly those situations that are covered by the generalized explanation.

In our setting, the non–operational background knowledge is the qualitative perception model, training examples would be sections of a melody harmonized with some specific chords, and the system would have to explain why each of the chords is a good harmonization. The outcome would be operational rules that would specify classes of musical situations where each of these chords may be used. However, the EBL technique is not directly applicable in our case, since it requires that the knowledge available to the learner be complete: every training instance must be fully explainable on the basis of this knowledge. This is clearly not the case with our perception model, which is too abstract and incomplete with respect to specific harmonization tasks.

For this reason, we have developed a new learning algorithm that can use all the knowledge that seems pertinent to the current situation, but does not depend on its completeness. Using this algorithm, the system learns faster and more intelligently: the perception model points to plausible partial explanations for given situations, and finer discriminating conditions are learned empirically by comparing various known situations.

The program learns in an interactive setting: We assume that a human teacher is present who presents sample harmonizations to the program and also critiques harmonizations of new pieces constructed by the program. The overall schematic structure of the learning scenario is sketched in Fig.3.

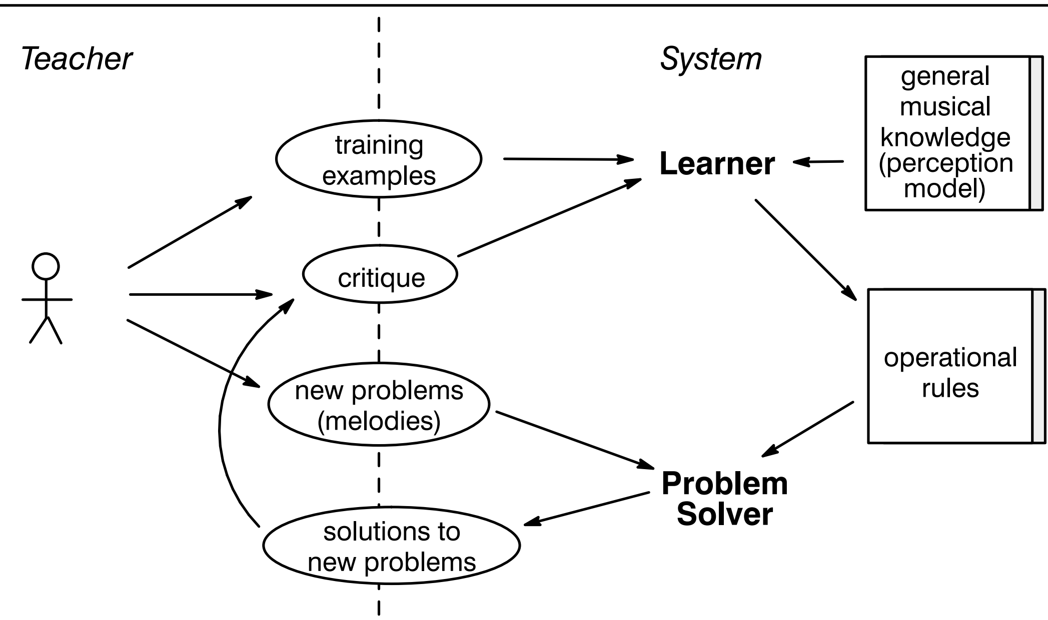


Fig.3: Learning scenario

The program itself consists of two subsystems: the Learner and the Problem Solver. The Learner analyzes harmonizations provided by the teacher and learns specific problem–solving rules from them. The Learner can also learn from mistakes committed by the Problem Solver. Finally, the Learner may also pose questions to the teacher; for instance, it may present some hypothesis to the teacher and ask him/her to confirm or reject it.

The Problem Solver uses the rules learned by the Learner to solve given harmonization problems. The search strategy employed by the Problem Solver is a left–to–right search with voting and backtracking.

The problem–solving algorithm:

Start with CurrentNote = first note of the melody and LastChord = NIL.

while the end of the melody has not been reached:

if there is some rule of type harmonic_change that matches the current situation
(and thus indicates that a new chord should be chosen to accompany CurrentNote)
then collect all rules of type good(Chord, CurrentNote) that apply in this situation;
sort the chords proposed by these rules according to the number of rules that propose them;
if there are chords proposed by some rules
then select the best of these chords, tentatively place it underneath CurrentNote, and go on to the next note;
else backtrack to the last note with open choices and try the next best chord there
else leave LastChord as a harmonization for CurrentNote and go on to the next note

The learning algorithm

The learning algorithm tries to use as much of the available musical knowledge as possible, while also paying due attention to the example harmonizations that are shown by the teacher. The algorithm is based on the notion of plausible explanations (cf. DeJong 1989). Basically, it proceeds as follows:

The Learner is given a *training example* by the teacher (where a training example is some melody with accompanying chords; one chord is designated as *focus*; this is the chord from which some lesson should be learned; the teacher also indicates whether this chord is indeed a good solution). Let us call this the *current example*.

If the current example was labelled as good by the teacher, the program first tries to *explain* it; that is, it tries to construct a plausible explanation, on the basis of its perception model, of why this designated chord might be considered good. Various heuristics are applied to assess the relative plausibility of such explanations, based on degree of match, proportionality of related values, etc. The leaves of the resulting explanation tree refer to surface features of the current example, that is, they point to operational features (features that can more or less directly be observed in a musical passage, or can be deduced in a straightforward way).

If the system succeeded in finding a consistent explanation, this explanation is presented to the teacher as a first hypothesis, and if s/he agrees, the explanation is generalized by turning constants to variables in an appropriate way. The leaves of this generalized explanation are then collected and conjoined to form the condition part of a new rule of type good(Chord, Note). Since the leaves represent operational conditions, the resulting rule will also be operational.

If no plausible explanation can be found, the example is simply stored. If some other, similar, examples are encountered later on, the system will try to find an empirical generalization that covers all these similar examples, and no conflicting ones, and will convert this empirical generalization into a rule. In this way, the system can also learn rules and concepts that are not explainable on the basis of the perception model, but represent stylistic conventions or idiosyncrasies of the teacher's harmonization style.

Rules of the second type – those with head harmonic_change(LastChord, CurrentNote), which specify when a change of harmony is called for – are learned in a similar way: the system knows, e.g., that people feel the need for a new harmony when either the last chord is perceived as highly unstable in the hypothesized local key, or when a high degree of certain kinds of tension has been building up. To learn a rule of type harmonic_change, then, requires verifying that some of these conditions may hold in the current situation. However, not all conventions of harmonic change can be explained in this way. Consequently, many of the harmonic_change rules have to be learned by empirical generalization.

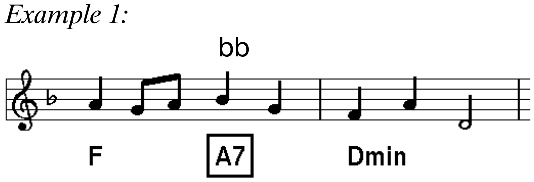
A final point is that learned rules are not static: when the system encounters a musical situation that 'almost' fits a learned rule, i.e., would only require minor modifications to the rule in order to match, these modifications are carried out and the rule is replaced with its generalized version. In this way, rules that are too specific (because of some inessential details of the particular example from which they were learned) will be gradually generalized to cover broader classes of situations. This stepwise generalization is itself an intricate process; it is based not only on a comparison of a learned rule with a new learning situation, but also on an extensive analysis of the explanation from which the rule was originally derived. For a detailed discussion of the heuristics used see (Widmer 1991a).

The length and complexity of this verbal description reflects the complexity of the algorithm itself. In order to give the reader at least some idea of what this all means, we will proceed with a relatively simple example.

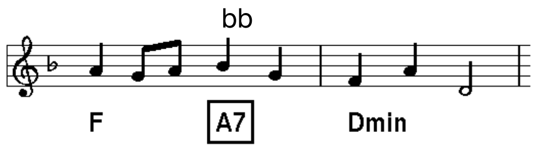
An example

The system is presented with the following musical situation (global key is F major):

Example 1:



The harmonization is given by the teacher, and the current focus of attention is the A7 chord in measure 1. Now the system sets out to find a plausible explanation of why A7 might be a good harmonization for the b–flat in the melody. (In a next step, it will also investigate the role of A7 as a harmonization for the subsequent g in the melody.) It finds an explanation that is based directly on the qualitative perception model (Fig.4). This time we list the explanation exactly as it is output by the program, in order to give a flavor of the real complexity of the system. The branches of explanations are annotated according to the different types of explanatory links (DED for deductive arguments, QPLUS for arguments based on Q+ proportionalities in the perception model, etc.), and each branch is also annotated with an assessed degree of plausibility. This differentiated information is very important for empirical generalization steps that may be necessary later on – see (Widmer 1991a).



type of explanatory link
assessed degree of plausibility
qualitative value of domain parameter (from qualitative model)

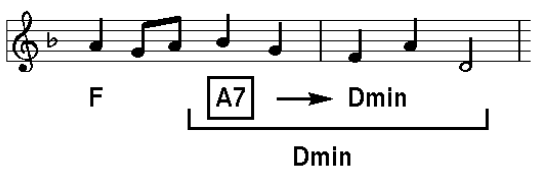
```
good(A7,bb)  <– DED (high) ––
  characterization(A7)  <– DED (extremely_high) ––
    chord_root(A7,a)  <– TRUE (extremely_high)
    chord_mode(A7,major)  <– TRUE (extremely_high)
    chord_type(A7,7)  <– TRUE (extremely_high)
    global_key(key(f,major))  <– TRUE (extremely_high)
  goodness(A7,bb,extremely_high)  <– QUAL (high) ––
   goodness(A7,bb,extremely_high)  <– QPLUS (high) ––
    interestingness(A7,bb,extremely_high)  <– QUAL (high) ––
     interestingness(A7,bb,extremely_high)  <– DEP (moderate) ––
      contrast(F,A7,moderate)  <– QUAL (extremely_high) ––
       contrast(F,A7,moderate)  <– QPLUS (extremely_high) ––
        previous_chord(A7,F)  <– TRUE (extremely_high)
        relative_chord_distance(F,A7,moderate)  <– DED (extremely_high) ––
         distance_on_circle_of_fifths(F,A7,4)  <– TRUE (extremely_high)
     interestingness(A7,bb,extremely_high)  <– QPLUS (high) ––
      tension_buildup(A7,bb,extremely_high)  <– QUAL (high) ––
       tension_buildup(A7,bb,extremely_high)  <– QPLUS (high) ––
        previous_chord(A7,F)  <– TRUE (extremely_high)
        local_tension_increase(F,A7,high)  <– DED (extremely_high) ––
          intrinsic_chord_tension(F,low)  <– DED (extremely_high) ––
           chord_type(F,triad)  <– TRUE (extremely_high)
          intrinsic_chord_tension(A7,high)  <– DED (extremely_high) ––
           chord_type(A7,7)  <– TRUE (extremely_high)
       tension_buildup(A7,bb,extremely_high)  <– QMINUS (extremely_high) ––
        relative_consonance(A7,bb,extremely_low)  <– DED (extremely_high) ––
         not_chord_contains_note(A7,bb)  <– TRUE (extremely_high)
         chord_scale(A7,scale(a,mixolydian))  <– TRUE (extremely_high)
         not_scale_contains_note(scale(a,mixolydian),bb)  <– TRUE (extremely_high)
       tension_buildup(A7,bb,extremely_high)  <– ADD.QPLUS (moderate) ––
        salience(bb,moderate)  <– DED (extremely_high) ––
         structural_salience(bb,moderate)  <– TRUE (extremely_high)
         metrical_strength(bb,moderate)  <– TRUE (extremely_high)
   goodness(A7,bb,extremely_high)  <– QPLUS (high) ––
    coherence(A7,bb,extremely_high)  <– QUAL (high) ––
     coherence(A7,bb,extremely_high)  <– QPLUS (high) ––
      harmonic_stability(A7,bb,extremely_high)  <– QUAL (high) ––
       harmonic_stability(A7,bb,extremely_high)  <– QPLUS (high) ––
        plausible_local_key(bb,key(d,minor))  <– TRUE (extremely_high)
        stability_of_chord_in_key(A7,key(d,minor),high)  <– DED (extremely_high) ––
          dominant_7_chord(key(d,minor),A7)  <– TRUE (extremely_high)
       harmonic_stability(A7,bb,extremely_high)  <– ADD.QPLUS (moderate) ––
        salience(bb,moderate)  <– DED (extremely_high) ––
         structural_salience(bb,moderate)  <– TRUE (extremely_high)
         metrical_strength(bb,moderate)  <– TRUE (extremely_high)
```

Fig.4: Training example and plausible explanation as displayed by the system

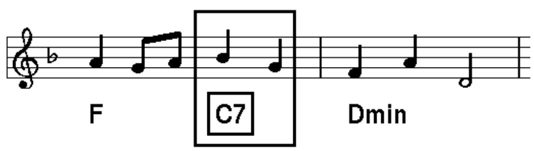
The explanation is presented to the teacher, who confirms that this does indeed make sense. The leaves of the explanation tree are collected and generalized, and duplicates are removed. The result is the following general, operational rule (learned rules are regular Prolog clauses; variables are capitalized; mnemonic variable names were substituted by the author; Prolog syntax was enhanced by explicitly adding the words IF and and for better legibility):

```
Rule 1:  good( Chord, Note)  IF
                 chord_root( Chord, Root)
             and chord_mode( Chord, major)
             and chord_type( Chord, 7)
             and global_key( Key)
             and previous_chord( Chord, PrevChord)
             and distance_on_circle_of_fifths( PrevChord, Chord, 4)
             and chord_type( PrevChord, triad)
             and not_chord_contains_note( Chord, Note)
             and chord_scale( Chord, Scale)
             and not_scale_contains_note( Scale, Note)
             and structural_salience( Note, moderate)
             and metrical_strength( Note, moderate)
             and plausible_local_key( Note, LocalKey)
             and dominant_7_chord( LocalKey, Chord)
```

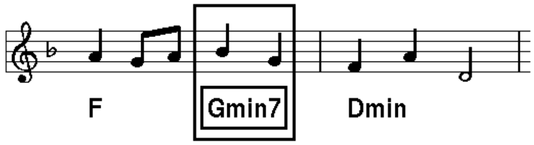
Note what has happened: The system has learned – from one particular harmonization provided by the teacher – that it is a good idea to precede/prepare a new local tonic (D minor, in our case) with an appropriate dominant seventh chord (A7), even if this dominant chord seems harmonically unstable with respect to the simultaneous portion of the melody (the b–flat is highly dissonant with the A7 chord, so the system could not verify relative_consonance(A7, b_flat, high), which is specified in the perception model as a plausible precondition for harmonic_stability). It has learned that in this case, harmonic stability of the chord within the implied local key, and the smoothness of progression to the subsequent local tonic afforded by the A7 chord, is more important than local harmonic consonance with the melody. That is,



is better (according to the teacher, who chose this harmonization) than, say,



or



And, what is more, the system has converted its experience with this particular learning situation into a general rule which will allow it to solve many similar problems in an analogous way, whether or not they share the particular surface features of this situation.

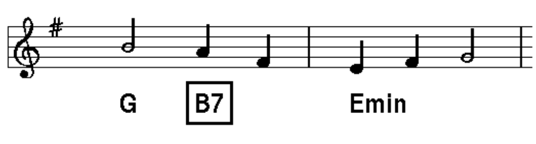
Note that, without its qualitative perception model, the system would be at a total loss. It would have to consider myriads of (to us, obviously non–sensical) hypotheses, without being able to determine which ones are plausible and which ones are not. For instance, it might just as well hypothesize that the A7 chord is a good harmonization because it is an interval of a major third up from its predecessor chord, which would lead to the absurd rule

```
good( Chord, Note) :–
        previous_chord( Chord, PrevChord),
        chord_root( Chord, Root),
        chord_root( PrevChord, Root1),
        is_interval( Root1, Root, maj3, up).
```

or that the A7 is good because its root forms a minor second with the note in the melody, which is equally absurd. There are many more such possibilities.

Finally, let us briefly indicate how the additional learning examples lead to empirical generalization. Let us assume that later on, the system encounters the following situation, where the B7 chord is classified as good by the teacher (global key is G major):

Example 2:



Examining the rule just learned, along with its underlying explanation (Fig. 4), the system notices that a slight generalization would suffice to make the explanation also fit example 2: the only relevant difference between the two situations (with respect to the explanation) is that in example 2, the chord in question (B7) is consonant with the note in the melody. Thus, by dropping the conditions not_chord_contains_note and not_scale_contains_note, and, by implication, dropping the entire branch relative_consonance(_, _, extremely_low), the above explanation structure can be generalized to describe both examples. Since there are still enough factors indicating that the solution is interesting (interestingness(_, _, extremely_high)), this has only a minor effect on the overall character of the explanation, so that this would seem a highly plausible generalization. The modified explanation is presented to the teacher, who will, of course, accept it as highly sensible. The resulting rule then describes a more general class of situations where a new local tonic is preceded by a corresponding dominant seventh chord, regardless of whether or not the dominant chord is consonant with the note in the melody:

```
Rule 1':  good( Chord, Note)  IF
                chord_root( Chord, Root)
          and   chord_mode( Chord, major)
          and   chord_type( Chord, 7)
          and   global_key( Key)
          and   previous_chord( Chord, PrevChord)
          and   distance_on_circle_of_fifths( PrevChord, Chord, 4)
          and   chord_type( PrevChord, triad)
          and   structural_salience( Note, moderate)
          and   metrical_strength( Note, moderate)
          and   plausible_local_key( Note, LocalKey)
          and   dominant_7_chord( LocalKey, Chord)
```

Implementation

The qualitative perception model and the program operating on it (Learner, Problem Solver) are fully implemented in Quintus Prolog on an Apollo workstation. The qualitative model itself consists of about 1000 lines of prolog; in addition, the underlying Lerdahl & Jackendoff model is another sizeable Prolog program. Prolog allows us to express the dependencies in the perception model in a declarative way; for instance, the fact that the stability of a chord in the currently perceived key contributes positively to perceived harmonic stability is represented directly as

```
q_plus( stability_of_chord_in_key( Chord, Key, S1),
      harmonic_stability( Chord, Note, S)) :-
   plausible_local_key( Note, Key).
```

Note that this Prolog clause consists of two parts: the head describes the proportionality relation itself, whereas the body (the part of the clause after the :– functor) specifies conditions for the proportionality to be applicable and establishes a context; in this case, the condition plausible_local_key(Note, Key) first establishes the local tonality that might be heard in the vicinity of the current note; this is the key with respect to which the proportionality will then be assessed. Since these statements are regular Prolog clauses, they can be backtracked through, so if there are several local tonalities that could plausibly be postulated, all of them will be considered in turn when searching for the best explanation.

Explanations are constructed by a Prolog meta–interpreter that can combine strict Prolog rules, qualitative proportionality statements, and determinations from the perception model. Learned rules are again regular Prolog clauses that can be directly executed by the problem solving component.

The complete system (Model + Learner + Problem Solver) comprises about 6000 lines of Prolog code. In addition, we have written a simple graphical music editor in C. It displays notes and chord symbols and allows the user to edit pieces graphically and to communicate with Learner and Problem Solver via a mouse–sensitive interface.

Despite the complexity of the model and the program, the system is rather fast: it takes at most 10 seconds to construct a complex explanation, and generalization of rules happens at the same speed. This indicates that it should be feasible to generalize this approach to more complex types of music and still get decent response times.

Conclusion

In summary, we believe that this article has made two important contributions: it has demonstrated the importance of *a priori* knowledge for an intelligent learning system and has presented an algorithm that maximally exploits general knowledge when learning new rules; this part of the research is independent of the particular area of application and is interesting to the field of Machine Learning in general. On the other hand, the basic musical knowledge we have given to the system was meant to be a general, plausible model of some aspects of music perception; we hope to have demonstrated a new approach to modelling musical knowledge at an abstract, yet still useful level.

But despite the success of our qualitative perception model in the learning experiments, we are well aware of the limitations of the model as it now stands. Too much is missing from the model (for instance, the real−time dimensions of listening, questions of the perception of form, dimensions like rhythm, timbre, etc.) for it to be considered a truthful picture of human music listening. The question is whether the approach itself is musicologically sensible, and if so, whether it can be generalized to more complex forms of music. Not surprisingly, our own response to these two questions would be to the affirmative. We do believe that important phenomena of perception and cognition can be modelled at an abstract level, and that this will be a fruitful way both for theoretical and practical musical systems.

On the theoretical side, projects of this kind should help shed more light on the cognitive aspects of music, and on the prerequisites for what might be called 'Musical Intelligence' (or, if nothing else, on our lack of knowledge and the need for more theoretical research). On the practical side, one can easily imagine how this type of research might lead to musical systems of practical utility. Take, for instance, the case of an intelligent composer's aid, a computer program that would interactively assist a composer in his or her work. If we can construct qualitative models of more complex types of music than considered here, such a computer program could gradually adapt to the composer's style and intentions, and would become more and more sophisticated, and thus more useful to the composer − think of a system like David Cope's EMI program (Cope 1987), but much more flexible, because less fixed in its built−in structures.

Our own immediate goals for further research, then, are to extend our perception model by including listening dimensions that have been neglected in the current version and to generalize the model to more complex types of music. In particular, we want to incorporate a notion of *time* in the model, which is so crucial in the actual process of listening. To this end, we are investigating the possibility of using alternative theories of music (instead of the Lerdahl & Jackendoff model) as a basis. Eugene Narmour's *Implication−Realization Theory* (Narmour 1977) appears as a promising candidate, for it is a much more plausible theory of how music might actually be constructed during the act of listening.

Acknowledgements

Comments by Ernst Buchberger, Christian Holzbaur, John Matiasek, Paolo Petta, and an anonymous referee greatly improved the quality of this paper. Thanks to Prof. Robert Trappl for his continuing support. This research was sponsored in part by the Austrian *Fonds zur Förderung der wissenschaftlichen Forschung*. Financial support for the Austrian Research Institute for Artificial Intelligence is provided by the Austrian Federal Ministry for Science and Research.

References

Bharucha, J. and P. Todd. 1989. "Modeling the Perception of Tonal Structure with Neural Nets." *Computer Music Journal* 13(4): 44–53.

Collins, A. and R. Michalski. 1989. "The Logic of Plausible Reasoning: A Core Theory." *Cognitive Science* 13: 1–49.

Cope, D. 1987. "An Expert System for Computer–assisted Composition." *Computer Music Journal* 11(4): 30–46.

DeJong, G. 1989. "Explanation–Based Learning with Plausible Inferencing." In K.Morik, ed. 1989. *Proceedings of the Fourth European Working Session on Learning (EWSL–89).* London: Pitman.

DeJong, G. and R. Mooney. 1986. "Explanation–Based Learning: An Alternative View." *Machine Learning* 1(2): 145–176.

de Kleer J. and J.S.Brown. 1984. "A Qualitative Physics Based on Confluences." *Artificial Intelligence* 24(1–3): 7–84.

Ebcioglu, K. 1988. "An Expert System for Harmonizing Four–part Chorales." *Computer Music Journal* 12(3): 43–51.

Forbus, K. D. 1984. "Qualitative Process Theory." *Artificial Intelligence* 24(1–3): 85–169.

Kuipers, B. J. 1986. "Qualitative Simulation." *Artificial Intelligence* 29(3): 289–338.

Laske, O. 1988. "Introduction to Cognitive Musicology." *Computer Music Journal* 12(1): 43–57.

Lerdahl, F. and R. Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge: The MIT Press.

Longuet–Higgins, H. C. 1988. "Perception of Melodies." In W. Richards, ed. 1988. *Natural Computation*. Cambridge: The MIT Press.

Michalski, R. 1983. "A Theory and Methodology of Inductive Learning." In R. Michalski, J. Carbonell, and T. Mitchell, eds. 1983. *Machine Learning: An Artificial Intelligence Approach.* Palo Alto: Tioga.

Mitchell, T., R. Keller, and S. Kedar–Cabelli. 1986. "Explanation–Based Generalization: A Unifying View." *Machine Learning* 1(1): 47–80.

Narmour, E. 1977. *Beyond Schenkerism*. Chicago: University of Chicago Press.

Sloboda, J. A. 1985. *The Musical Mind: The Cognitive Psychology of Music.* Oxford: Oxford University Press.

Widmer, G. 1990. *A Qualitative Model of the Perception of Melodies and Harmonizations*. Report TR–90–12. Vienna: Austrian Research Institute for Artificial Intelligence.

Widmer, G. 1991a. ”Using Plausible Explanations to Bias Empirical Generalization in Weak Theory Domains.” In Y.Kodratoff, ed. 1991. *Proceedings of the 5th European Working Session on Learning (EWSL−91)*. Berlin: Springer Verlag.

Widmer, G. 1991b. ”Analytical and Empirical Generalization with Qualitative Domain Theories.” *8th International Workshop on Machine Learning,* Northwestern University.

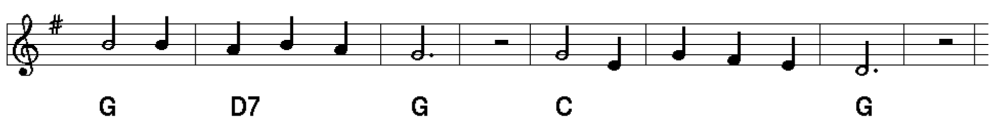
Widmer, G. 1991c. ”A Knowledge Intensive Approach to Machine Learning in Music.” In M. Balaban, K. Ebcioglu, and O. Laske, eds. 1991. *Musical Intelligence*. Menlo Park: The AAAI Press. (in press)

Appendix: An Experiment

We present here a description of a small experiment that gives a flavor of how fast and effective the learning process is. The entire experiment took no more than twenty minutes. The basic setup was as follows: we prepared two training pieces, complete with given harmonization, and a test piece (melody only) which the system should try to harmonize at different stages of the learning session. The three pieces are the beginnings, respectively, of three well−known Austrian Christmas carols:

Training piece 1: ”Kling Glöckchen”

Training piece 2: ”Leise rieselt der Schnee”



Test piece: ”Stille Nacht, heilige Nacht” (”Silent Night, Holy Night”)

At the beginning of the experiment, the system starts from scratch, that is, without any specific harmonization rules. If presented with the test piece at this stage, it will fail to find any solution. In fact, it will not even try, because there are no rules to suggest possible chords and places for chord changes, and blindly trying all possible combinations of chords and places to find a harmonization that the perception model classifies as good would be a hopelessly expensive affair.

Now the system is presented with training piece 1 (”Kling Glöckchen”). It is told that all the chords in this harmonization are good. The system finds plausible explanations for the acceptability of all of these chords, and it also succeeds in verifying that the points in the piece where a new chord enters can reasonably be interpreted as points of perceived harmonic change. The result of this analysis is a set of seven specific rules (three of them concerning conditions for a chord change to occur, and four suggesting possible chords under various circumstances). An example of each of these types of learned rule follows (the rules were slightly edited and variables substituted with mnemonic names in order to improve readability):

Rule 2:

```
harmonic_change( LastChord, CurrentNote)  IF
          not_chord_contains_note( LastChord, CurrentNote)
    and   chord_scale( LastChord, Scale)
    and   scale_contains_note( Scale, CurrentNote)
    and   metrical_strength( CurrentNote, extremely_high)
    and   structural_salience( CurrentNote, high).
```

Translation: ”If the last chord was LastChord and the current note is CurrentNote, it may be a good idea to switch to a new chord at the point marked by CurrentNote if the following conditions hold:

LastChord does not contain (i.e., is not strictly consonant with) CurrentNote,

the scale implied by LastChord is Scale,

this Scale contains CurrentNote (so LastChord would be weakly consonant with CurrentNote),

CurrentNote is in an extremely strong metrical position,

and CurrentNote is also highly structurally salient (in terms of time−span reduction).”

One may note that the conditions calling for weak consonance between the last chord and the current note (conditions 2 and 3) are really too strong: if the previous chord were not even weakly consonant with the current note, then the system would have all the more reason to switch to a new harmony. When the system encounters such a training situation, the rule will automatically be generalized by dropping these overly restrictive conditions.

Rule 4 below is an example of the second type of rule learned from training piece 1:

Rule 4:

```
good( Chord, CurrentNote)  IF
          chord_root( Chord, ChordRoot)
    and   chord_mode( Chord, major)
    and   chord_type( Chord, triad)
    and   global_key(  Key)
    and   tonic_chord( Key, Chord)
    and   chord_contains_note( Chord, CurrentNote)
```

```
and  plausible_local_key( Note, Key)
and  structural_salience( Note, extremely_high).
```

Translation: ”When trying to find a harmonization for note CurrentNote, any chord Chord that satisfies the following conditions will be a good candidate:

Chord is a major triad with root ChordRoot (conditions 1 through 3),

the global key of the piece is Key,

Chord is the tonic chord of Key,

Chord is highly consonant with (i.e., contains) CurrentNote,

Key is also a plausible local key in the surroundings of CurrentNote,

and CurrentNote is of extremely high structural salience.”

(This rule was learned from the first note−chord pair (A − D major) in training piece 1, where note A figures very prominently in the time−span reduction of the melody.)

At this stage we tell the system to harmonize the test piece (”Stille Nacht”), using the rules learned so far. Since, up to this point, the system has only gained experience with tonic and dominant seventh chords, it produces a solution involving just these two chord types:

(Solution found for test piece after training with piece 1)

Now the system is presented with the second training piece (”Leise rieselt der Schnee”), with the given harmonization classified as good by the teacher. Analysis of this piece prompts the system to learn five more rules, the most important of which derive from the C major chord in measure 5. Now the test piece is again presented for harmonization, and with the additional experience acquired from piece 2, a much better solution is found:

(Solution found for test piece after training with pieces 1 and 2)

This short experiment demonstrates how extensive analysis of given musical material with the help of general knowledge about musical effects enables a computer to acquire effective problem solving knowledge from a small number of specific experiences. The two training pieces used in the experiment dealt with only three types of chords (I, IV, V^7). But given a few more training pieces, the program learns more complex concepts like switching to a new local tonic, preparing this new tonic with dominant and subdominant chords, etc. (see also the example in the article).

We have also performed experiments with the beginnings of several Mozart piano sonatas. These experiments were specifically designed to measure the impact of the presence of musical background knowledge on incremental empirical generalization of rules. The arguments and results are presumably of interest only to specialists in Machine Learning; they are reported in (Widmer 1991a, b).