A MARKOV MODEL FOR CHORALE HARMONIZATION

Kaan M. Biyikoglu
Middle East Technical University, Informatics Institute, Ankara, Turkey

ABSTRACT

The research presented in this paper aims to study the syntax of harmony and the interaction of harmony and melody by building a Markov model for harmonization. The Markov model described in the present article was trained by the output of a harmonic analysis module operated on a collection of chorales, and was used to compute suitable chord progressions for novel melodies. After the calculation of a progression, four-part harmonizations were generated by matching input melody and calculated progression with the voicing patterns found in the chorales.

The paper discusses how Markov models can be used in harmonization tasks, various pitfalls of the Markovian approach, some strategies for improved performance, and also the effect of two pattern matching strategies on the outputs.

1. BACKGROUND

Harmonization has been a popular topic for computational studies of music (Ebcioglu, 1992; Hild, 1991; Cope, 1999; Phon-Amnuaisuk et al., 1999; Phon-Amnuaisuk et al., 2002), and various researchers have investigated the possibility of studying harmony by using statistical models (Thom, 1995; Ponsford et al., 1999). Harmonization task can be decomposed into two interacting sub-processes, one of which is the segmentation of melody according to an implied harmonic rhythm, and the other is the assignment of harmonic functions for each segment in order to satisfy at least the following conditions: (a) harmonic function of a segment shall be harmonious with the pitch events found in the segment; (b) the overall harmonic progression shall be satisfactory.

Markov models provide straightforward probabilistic estimates for the satisfactoriness of chord progressions. The intrinsic assumption of Markov models is that, given a history of events, the probability of a next event depends *only* on a fixed number of previous events. Making such an assumption enables us to define a measure for harmonic expectation subsequent to a series of previous harmonies in terms of probability values.

The crucial limitation associated with Markov models is that their descriptive power is equal to that of regular languages. A musical implication of this observation is that while Markov chains can effectively compute smooth modulations from one key to another, the common notion of returning to the home key after a series of modulations is beyond their capability (Chomsky, 1957).

Nevertheless, some limitations of the chorale form facilitate the construction of a Markov model of harmony by using chorales: chorales are short pieces usually having the simple *AAB* form; they do not possess structurally *deep* modulations; they have relatively monotonous rhythmic organization; one can find a relatively large number of chorales of a given style of music to form a database for analysis; and – most importantly – segmentation is not a big problem for two reasons: (1) the harmonic rhythm is mostly regular with harmonic changes occurring per beat, and (2) chorales – specifically Bach chorales – make up an annotated corpus, since the phrases are explicitly marked with the fermata signs on the score.

2. THE MODEL

A Markov model depends on the assumption that the probability of a next event is conditional only on a finite number of previous events. Let $w_1 \ldots w_t$ be a sequence of random variables taking values from a finite alphabet $\sum=\{c_1 \ldots c_n\}$ (e.g. a set of chords). Then the *Markov property* can be stated as

$$P(w_{t+1} = c_j \mid w_1 \ldots w_t) = P(w_{t+1} = c_j \mid w_{t-k+1} \ldots w_t)$$

where k is the *order* of the Markov process and it determines the length of the fixed window of histories used in calculations. For k=2 we have a 2nd order Markov process, such that

$$P(w_{t+1} = c_j \mid w_1 \ldots w_t) = P(w_{t+1} = c_j \mid w_{t-1} w_t)$$

Making the Markov assumption simplifies the calculation of the probability of a series of random variables. Assuming a 2nd order Markov model, the probability of a sequence of random events can be calculated as

$$\begin{aligned} P(w_1 \ldots w_t) &= P(w_1) \lozenge P(w_2 \mid w_1) \lozenge P(w_3 \mid w_1 w_2) \cdots P(w_t \mid w_1 \ldots w_{t-1}) \\ &= P(w_1) \lozenge P(w_2 \mid w_1) \lozenge P(w_3 \mid w_1 w_2) \cdots P(w_t \mid w_{t-2} w_{t-1}) \end{aligned}$$

The probability measures such as $P(w_t=c_k|w_{t-2}=c_i, \ w_{t-1}=c_j)$ are called the *transition probabilities*, and they designate the probability that the sequence $(c_i \ c_k)$ is followed c_k. In a Markov model of harmony, such transition probabilities indicate the probability that a particular harmonic event follows two given previous harmonic events.

Usually the transition probabilities are estimated by taking the *maximum likelihood estimates* from a training corpus as

$$P_{MLE}(w_t = c_k \mid w_{t-2} = c_i, w_{t-1} = c_j) = \frac{C(c_i c_j c_k)}{C(c_i c_j)}$$

where $C(c_i \ c_j \ c_k)$ and $C(c_i \ c_j)$ are the number of occurrences of the strings $c_i \ c_j \ c_k$ and $c_i \ c_j$ in the training corpus (Manning and Schütze, 1999, p.198). Once the transition probabilities are estimated from a training corpus, they can be used to estimate the probabilities of generated progressions.

ISBN 3-931852-67-9
ISSN 1617-6847

2.1. Representing Time and Meter in Markov Models

Markov models of harmony are intrinsically problematic to come up with an adequate representation of time in music, since they operate in a sequential manner, whereas time is hierarchical in music. Lerdahl and Jackendoff consider time-span reduction as a notational variant of prosodic structure (1983, p. 314), and Horton (2001) provides some examples in which the harmonic constituent formation changes when a progression is presented in different metric settings. Such properties of meter restrain the use of finite state techniques even more, since the concept of constituency is mostly alien to finite state models.

The usual ad-hoc solution for this problem – imported from corpus linguistics – is to use annotated strings for training. Annotation is performed by inserting additional symbols into the strings of chord progressions obtained from harmonic analysis in order to indicate the positions of bar lines and the locations of phrase boundaries.

While chorales usually possess a regular harmonic rhythmic with harmonies usually changing per beat, there are occasional situations in which a harmony has longer duration. The discrete nature of Markov models allow there different encoding strategies to cope with temporal deviations:

1. *Use fixed sampling interval*: A fixed sampling interval is selected (e.g. quarter note duration), and each symbol in the strings of training corpus correspond to the harmony for a definite quarter note segment. The longer durations are represented as repetitions in the strings. In this approach the probabilities of repeating events are calculated by using non-zero terms such as $P(a|aa)$ or $P(b|cb)$. A side affect of this strategy is that, whenever there is repeating symbols, there is also a loss of information about the previous events due to the usage of fixed number of past events in Markov models.

2. *Generate new symbols:* Additional symbols are provided in order to reflect the temporal variation among events. For example, there are distinct symbols, say c^1 and c^2, in the alphabet for C-major chords having quarter-note duration, and C-major chords having half-note duration. Hence there are separate transition probability measures such as $P(c^1|ab)$ and $P(c^2|ab)$ which reflect the durational variation. This strategy was utilized by Ponsford et al. (1999).

3. *Retreat:* Do not represent duration in the training strings. The obtained transition probabilities more securely reflect changes in harmonic functions, but the information about duration is lost.

All of the above approaches have their shortcomings; the 3rd approach leaves out the duration information; in the 1st first approach we loose information about the past events; both 1st and 2nd approaches involve the calculation of some probability measures whose relevance to expectation can be questioned in cognitive or music theoretic grounds, although they reflect some quantitative fact about the training corpus. Note that the 2nd strategy requires a Markov model with larger alphabet, and that each strategy necessitates a distinct way of preprocessing of the training strings.

2.2. Building the Markov Model of Harmony

The alphabet of the actual model consisted of chord symbols (major, minor, diminished and augmented triads, as well as major, minor, dominant, half-diminished, and diminished sevenths built in all 12 pitches), and four annotation symbols (*start, end, bar, phrase*). The phrases were determined according to the location of the fermata signs on the score. The annotated progression strings were the outputs of a harmonic analysis program operated on a corpus of Bach chorales, and they were all transposed to the same key prior to training. Different models were built for chorales having major and minor keys. The transition probabilities were estimated by using maximum likelihood estimates.

2.3. Using the Markov Model for Harmonization

Prior to the harmonization, input melodies are first segmented with respect to a desired harmonic rhythm (usually regular, segment duration being quarter-note duration), and then sets of chords that can be used are generated for each segment according to the following assumption:

Assumption 1: Let M be a melody, and $m_1m_2...m_T$ be its segmentation. Given a set of available chords $\sum$, we can calculate for each segment m_i of M the corresponding set of chords $\sum_i$ such that $\sum_i \subseteq \sum$ and a chord $c \in \sum$ is in $\sum_i$ if and only if at least one pitch class of c is present in m_i.

The following assumption states that the probability of usability of a chord for a segment conditional on the segment does only depend whether the chord is in the set of chords as given by the first assumption.

Assumption 2: Given a segment m_i and a corresponding set of candidate chords $\sum_i$, each chord in $\sum_i$ has equal probability of harmonizing m_i.

Given a segmentation of a melody M, as $m_1...m_T$, after the generation of the corresponding vector of sets $\sum_1 \dots \sum_T$, the above assumptions reduce the harmonization problem to a problem of finding a path that passes through the sets $\sum_i$ which has relatively high probability. Whenever there is information about the metrical organization of the melody to be harmonized and the location of cadences, the quality of harmonization can be improved by inserting singleton sets consisting of annotation symbols into appropriate positions of the vector $\sum_1 \dots \sum_T$. For a given melody, one can calculate a progression corresponding to a path with relatively a high probability by running the Markov model stochastically. Moreover, the most probable path according to the model that passes through the sets $\sum_i$ can also be obtained by using the Viterbi algorithm (Manning and Schütze, 1999; Biyikoglu, 2002).

ISBN 3-931852-67-9
ISSN 1617-6847

The process described above can occasionally generate problematic chord progressions. One such example is provided in Figure 1 which illustrates two harmonic violations. The quarter note with the pitch *A4* just prior to the first cadence is harmonized with a B-7th chord and the pitch of the melody here acts as a chordal 7th. Unless the next chord is also a 7th, in Bach's period the chordal 7th was usually resolved to a note which is one scale step below, but in this case the melody moves upwards to *B4*. The second problem occurs just at the beginning of the second phrase which is harmonized with a *D major* triad. Using D major chord at this location causes the *G4* of the melody to act as an appoggiatura at the beginning of the second phrase. Such an appoggiatura just at the beginning of a phrase is not common in chorale harmonizations of the period.

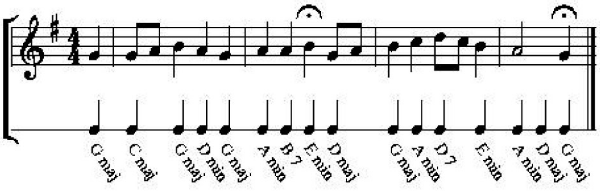

Figure 1: A problematic chord progression generated by the model for the first two phrases of the chorale *Nun freut euch, lieben Christen, g'mein*.

The above discussion demonstrates that the chord selection criteria should be refined by means of a more knowledge-rich inference mechanism which takes into account harmonic resolutions and non-harmonic tones. In the current implementation of the model, 7th chords are not allowed if their resolution is prevented by the melody, and appoggiaturas are not allowed at the beginning of phrases.

2.4. Voice Leading

Cope (1999) reports a strategy he implemented in EMI according to which the segments from a database of chorales, each having one-beat duration, are combined together for the composition of new chorales. The combination of two segments was possible if the initial pitches of one segment were equal to the destination pitches of another. In this work, the same strategy was used to generate four-part textures for the melodies whose chord progressions were determined. In the model, the constraint of having equal destination pitches were relaxed such that the combination was possible if the pitch-class intervals among voices produced by the combination were equal to the pitch-class intervals among the voices of the former segment and its destination. Moreover the combinations had to produce both the melody to be harmonized, and the harmony generated by the model.

Unfortunately a database of segments obtained from 170 Bach chorales was not enough, and only partial four-part texture could be generated for given chorale melodies. Such a harmonization of the chorale melody given in Figure 1 is displayed in Figure 2, which incorporate a chord progression different than the one given in Figure 2, and which was calculated by the Viterbi algorithm using a 2nd order Markov model.

Figure 2: A four-part harmonization of the melody given in Figure 1.

Another example is provided in Figure 3. The chord progression for the harmonization of Figure 3 was generated stochastically by a 3rd order Markov model. The harmonization is partial, that is, the mentioned voice-leading method could not generate a complete four-part texture for the given chorale melody and harmony.

Figure 3: A four-part harmonization of the first 5 phrases of the chorale *Herr Jesu Christ, du hast bereit*.

In order to overcome the problem of data sparseness, the combination constraints had to be relaxed, and this was tried by the incorporation of rules of voice-leading found in harmony textbooks (Gauldin, 1997). The database of chorale segments was retained, but this time, in addition to preserving the melody and harmony, segments had to obey the rules of voice-leading in the combinations (rules related with voice-crossing and overlap are not implemented, since there are enormous exceptions to these rules in Bach chorales). An exemplar to the voice-leadings generated with this approach is presented in Figure 4.

ISBN 3-931852-67-9
ISSN 1617-6847

Figure 4: A four-part harmonization of the melody given in Figure 1 for which the combination is guided by the rules of voice-leading.

Albeit the rules are not violated, the voices in the harmonization given in Figure 4 lack good continuation since most of the dissonances are unresolved. In order to obtain satisfactory results, the four-part texture generated by recombination and rules of voice-leading should be constrained to produce smooth inner voices.

3. DISCUSSION

Although Markov models are not capable of representing hierarchical relations of harmonic functions, they can generate smooth chord transitions, and in the presence of information about meter and cadence locations, are capable of generating acceptable chord progressions for given melodies. It should be noted that the use of annotation symbols for cadence positions in training and generation allow the model to generalize which chords are mostly used in cadences, and hence this important hierarchical information is to some extent captured by the model.

A subjective observation is that the progressions generated by the Viterbi algorithm, which are the most probable ones according to the model, tend to be conservative. Most of the time these progressions revolve around tonic and dominant, and are less adventurous compared with the progressions generated randomly.

The most important deficit of the present work is the lack of a good voice-leading component. The limitations of the recombination process prevent the generation of voice-leadings in many situations. Therefore one future work is to design a better voice-leading component, presumably a component which can infer the required characteristics from a corpus of chorales.

Another future work is to develop a Markov model having an alphabet consisting of intervals of progressions instead of chords themselves. Thom (1995) reports that using interval encoding instead of absolute encoding resulted in slight improvement in chord prediction task. Using interval encoding is also expected to allow more freedom in the generated progressions.

4. REFERENCES

1. Biyikoglu, K. M. (2002). A Markov Model for Chorale Harmonization in the Style of J.S. Bach. M. S. thesis, Department of Cognitive Science, Middle East Technical University.

2. Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

3. Cope, D. (1999). One approach to musical intelligence. *IEEE Intelligent Systems 14(3):21-25*.

4. Ebcioglu, K. (1992). An expert system for harmonizing chorales. In M. Balaban, K. Ebcioglu, and O. Laske (eds), *Understanding Music with AI: Perspectives on Music Cognition* (pp. 294-333). AAAI Press / MIT Press, Menlo Park, California.

5. Gauldin, R. (1997). *Harmonic Practice in Tonal Music*. W. W. Norton & Company, New York.

6. Hild, H., Feulner, J., and Menzel, W. (1991). HARMONET: A neural net for harmonizing chorals in the style of J.S. Bach. In J. E. Moody, S. J. Hanson, R. P. Lippmann (eds) *Advances in Neural Information Processing* (pp. 267-274), Morgan Kaufmann.

7. Horton, T. (2001). The compositionality of tonal structures: A generative approach to the notion of musical meaning. *Musicae Scientiae, V(2):131-159.*

8. Lerdahl F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. MIT Press, Cambridge, Massachusetts.

9. Manning, C. D. and Schütze, H. (1999). *Foundations of Statistical Natural Language Processing.* MIT Press, Cambridge, Massachusetts.

10. Phon-Amnuaisuk P., Smaill, A., and Wiggins, G. (2002). A computational model for chorale harmonization in the style of J.S. Bach. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (eds), *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 219-222), Sydney.

11. Phon-Amnuaisuk, P., Tuson, A., and Wiggins, G. (1999). Evolving Musical Harmonisation. In *Proceedings of ICANNGA'99*. Portonož, Slovenia.

12. Ponsford, D., Wiggins, G., and Mellish, C. (1999). Statistical learning of harmonic movement. *Research Paper 847*. University of Edinburgh, Division of Informatics, School of Artificial Intelligence.

13. Thom, B. (1995). Predicting chords in jazz: The good, the bad, and the ugly. In *IJCAI Workshop on AI and Music*. International Joint Conference on Artificial Intelligence IJCAI-95, Montreal.

ISBN 3-931852-67-9
ISSN 1617-6847