

Probabilistic harmonization with fixed intermediate chord constraints

Maximos Kaliakatsos–Papakostas
School of Music Studies,
Aristotle University of Thessaloniki
max@mus.auth.gr

Emilios Cambouropoulos
School of Music Studies,
Aristotle University of Thessaloniki
emilios@mus.auth.gr

ABSTRACT

During the last decades, several methodologies have been proposed for the harmonization of a given melody with algorithmic means. Among the most successful of them are methodologies that incorporate probabilistic mechanisms and statistical learning, since they have the ability to generate harmonies that statistically adhere to the harmonic characteristics of the idiom that the training pieces belong to. The current paper discusses the utilization of a well-studied probabilistic methodology, the hidden Markov model (HMM), in combination with additional constraints that incorporate intermediate fixed-chord constraints. This work is motivated by the fact that some parts of a phrase (like the cadence) are characteristic about the phrase’s idiomatic identity, as also supported by many works in the automated melodic harmonization literature. Among the targets of the ongoing research in the context of the COINVENT project, a part of which is the content of this paper, is to define and isolate such important parts and include them as constraints in a probabilistic harmonization methodology. To this end, the constrained HMM (CHMM) is developed, harnessed with the novel general chord type (GCT) representation, while the study focusses on examples that highlight the diversity that constraints introduce in harmonizations.

1. INTRODUCTION

Automated melodic harmonization discusses the assignment of harmonic material on the notes of a given melody. The harmonic material is described by chord symbols, while the harmonization is completed if the voicing of the notes that constitute the chords, as well as the voice leading between the notes of successive chords, is defined. The common approach to test an automatic harmonization system is to utilize it for harmonizing melodies that pertain to a musical idiom with harmonic structure that adheres to strict rules. To this end, some pioneering methodologies that were developed for melodic harmonization, incorporated human expert knowledge encoded in the form of rules, leading to expert systems [1] that could generate harmonizations with explicit stylistic orientation towards the mu-

sical idiom that these rules referred to. For a review in the rule-based systems the reader is referred to [2]. A similar approach to the rule-based methodologies is the one followed by systems that utilize genetic algorithms (GA), like the ones shortly reviewed in the recent paper [3] and in [4]. The similarity between these two approaches is that both rely on a set of harmonic rules intended for a specific musical idiom; in the case of the GAs, the employed fitness function quantifies such rules.

However, the rule-based spectrum of methods has a major drawback when discussing melodic harmonization in many different idioms: the encoding of rules that describe an idiom is not always a realizable task, since several idioms abound in contradicting and complex interrelations between harmonic elements. To this end, the formulation of *probabilistic* techniques and *statistical learning* has been proposed. Among many proposed methodologies, most of which are discussed in Section 2, Bayesian networks [5] and prediction by partial matching [6] have been utilized to construct the bass, tenor and alto voices below a given soprano voice, hidden Markov models (HMMs) for constructing chord sequences for a given melody [7] and probabilistic graphical models for relative tasks [8].

The approach to harmonization that is pursued in this paper, pertains to the wider research context of the COINVENT project, according to which the study of automatic melodic harmonization includes the blending of harmonic concepts among many musical idioms, to produce novel harmonic concepts. To this end, the exploration of harmonically meaningful chords within musical phrase are considered as distinctively important parts of an idiom. Such parts will be subsequently used as independent blend-able entities, allowing the mechanism of conceptual blending to produce harmonic “checkpoints” that comprise harmonic characteristic from multiple harmonic idioms. An example of structurally important parts are the chords in cadences, as discussed in the literature review presented in Section 2. However, the presented approach generalizes the notion of “important” chords to a methodology that allows the insertion of fixed-chord constraints in predefined positions of a phrase.

In this context, the potential of utilizing a well-studied probabilistic technique, namely the hidden Markov model (HMM), is promising, since this technique has yielded outstanding results in capturing the stylistic orientation of the idiom that is composed by the training pieces. To this end, the paper at hand proposes to tackle melodic harmonization through a mixing of methodologies: HMM with fixed-constrained intermediate chords. Thereby, the proposed

methodology is utilized to produce probabilistic melodic harmonizations that adhere to several fixed–chord constraints in intermediate checkpoints of the melody, as discussed in [9]. Harmonization with fixed checkpoints is considered a crucial component of the presented work, since it enables the prior definition of important chords in intermediate positions of the melody to be harmonized. The intermediate or “*anchor*” chords of a phrase are considered to be given either from an algorithmic process in a hierarchical level above the one discussed by the “chord progression” level – where chord transitions are defined by the the discussed HMM variation – or by a human user. However, it is beyond the scope of this paper to discuss potential algorithms for intermediate chord selection. Therefore, the experimental results mainly encompass examples where the fixed–chord constraints are provided either by a human experts, or by the chords utilized in the genuine composition of the harmonized melody.

An additional fundamental concern of the proposed harmonization approach is the idiom–independency in the chord symbols, chord relations and melodic considerations. This concern is addressed by utilizing the general chord type (GCT) representation, which is briefly discussed in Section 3.1. The presented algorithm acts on a certain level of the harmonic hierarchy, namely the *phrase* level. Thereby, given some “anchor” chords that remain fixed in a phrase, the aim of the algorithm is to select a “*proper*” chord sequences that connect the intermediate parts of the fixed chords, under the conditions introduced by the melodic material to be harmonized. The evaluation of the algorithm incorporates a comparison between the proposed constrained HMM (CHMM) and a “typical” HMM, which incorporates prior probabilities for the beginning and ending chords. The results indicate that CHMMs produce harmonizations that might be completely different to the ones produced by HMMs, depending on the imposed constraints. The results are reported on phrases of a set of J. S. Bach chorales, since they comprise an unofficial “benchmark” dataset for melodic harmonization methodologies.

2. PREVIOUS WORK AND MOTIVATION

Hidden Markov models (HMMs) have been extensively used for the automatic harmonization of a given melody, since their formalization describes the targeted task very well: given a sequence of observed notes (melody), find the most probable (hidden) sequence of chords that is compatible with the observations, according also to a chord transition matrix. In several studies of HMM–based melodic harmonization methodologies, a straightforward distinction is made on the role that some chords play to the composition – mainly the cadence of the phrase. For instance, the cadences of produced harmonizations by the HMM developed in [10] were utilized to evaluate the system’s performance, by comparing the cadence patterns that were produced by the system to the ones observed in the dataset.

Several HMM approaches discuss the utilization of some methodological tools to amplify the role of the cadence in the harmonization process. For instance, in [11] and [12] a backwards propagation of the HMM methodology is pro-

posed, i.e. by examining the prior probabilities of the final chord given the final melodic note. The Markov decision process followed in [13], rewards the authentic cadences thus providing higher probabilities to chord sequences that end with an authentic cadence. In [14] the phrases are divided in tonic, subdominant, dominant and parallel tonic chords, allowing a trained HMM to acknowledge the positions of cadences, however the selection of chords is performed through a rule–based process.

A commercial application utilizing HMM for melodic harmonic is *mySong* [15], which receives the melody by the singing voice of the user, extracts the pitches of the melody and employes an HMM algorithm to provide chords for the melody. The approach followed therein is equivalent to the one described in Section 3.1 (and in Equation 1), which is also used as a starting point towards the formalization of the BCHMM. There are also some additional features offered as harmonization options, like the “jazz” and the “happy” factors, but their description is beyond the scope this paper. According to the HMM approach utilized by *mySong*, prior probabilities are considered not only for the beginning chord of a piece, but also for the ending one, a fact that further biases the choice of solutions towards ones that incorporate first and final chords that are more often met in the training dataset.

The approach presented in this paper is motivated by the research in the aforementioned works, but it is completely on a fundamental aspect: it allows the *deterministic* (not probabilistic) insertion of chords at any place in the chord sequence. The automatic harmonization research area of the COINVENT project, a part of which is the work presented in the paper at hand, discusses the definition and extraction of several harmonic style (idiom) indicators, which will consequently be utilized to build composite harmonization that comprise characteristics from various idioms. To this end, the isolation of the harmony in “strategic” harmonic positions (e.g. the cadence, the beginning or intermediate parts of a phrase) is expected to contribute to the project’s perspective.

3. INTERMEDIATELY–CONSTRAINED PROBABILISTIC HARMONIZATION

The aim of the proposed methodology is to allow the probabilistic harmonization, while allowing prior determination of intermediate chords (also named as *checkpoints* in the literature [9]). The intermediate chords may either be specified by an algorithmic process that determines music structure on a higher hierarchical level, or may be directly inserted by a human annotator. Some examples of algorithm classes on higher hierarchical levels that could be utilized for providing intermediate anchor chords are rule-based approaches, generative grammars, or even Markov models trained with chords on a sparser time scale (e.g. the beginning, the middle and the final chord of phrases). Additionally, the fact that direct human intervention is enabled, allows the presented methodology to be the backbone of a melodic harmonization assistant, which allows its user to specify a harmonic “spinal chord” of anchor chords that are afterwards connected by chord sequences

that give aesthetic reference to a learned idiom.

An abstract example of a melodic harmonization process that incorporates some fixed anchor points is demonstrated in Table 1. Therein, a melodic line denoted by m_i $i \in 1, 2, \dots, 8$ (supposed with length 8) is harmonized with some given intermediate chords as constraints, namely I_i , $i \in 1, 2, 3$. The intermediate chords have been applied to specific notes of the melody, i.e. I_1 on m_1 , I_2 on m_5 and I_3 on m_8 . The first and final notes are harmonized with fixed chords for demonstration purposes, either one or both of them could be harmonized automatically by the variation of the HMM variation discussed in this papers. After the intermediate fixed chords have been defined, the *boundary-constrained* variation of the HMM (BCHMM) is utilized for each of the successive parts that begin and/or end with a fixed chord. In the case where the beginning and ending chords of the phrase are not fixed, the boundary constraints apply only on the fixed edge; the non-fixed edge is harmonized by utilizing the typical probabilistic HMM boundary condition, as discussed in the next paragraphs. For the example in Table 1, the BCHMM algorithm is applied twice, once of each pair of consecutive anchor chords – namely BCHMM¹ for connecting I_1 with I_2 and BCHMM² for connecting I_2 with I_3 .

Table 1. Abstract example of the proposed harmonization algorithm. On top (m_i) the melody notes to be harmonized are illustrated, below (I_i) the chord that are given as constraints and in the bottom (C_i^j) the chords produced by the j -th BCHMM method application.

mel.	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
con.	I_1				I_2			I_3
		C_1^1	C_2^1	C_3^1		C_1^2	C_2^2	
	BCHMM ¹				BCHMM ²			

The presented algorithm discusses only the level of chord labeling, i.e. its goal is to attribute a chord symbol – expressed as a GCT structure – disregarding information about harmonic rhythm and voice leading. Harmonic rhythm is a crucial matter that defines a vital part of a harmonization’s character, however, within the context of the prototypical evaluation of the proposed method, a chord is considered to accompany every note of the melody; a similar approach has often been endorsed in past research. Similarly, voice leading is also an important aspect of harmonization, while one could arguably consider that voice leading is sometimes fundamental in a sense that the movement of each separate voice defines the final vertical shape of the harmonic blocks. Such an argument would be more evidently valid for some idioms, like the polyphonic songs of Epirus, where the existence of steady-pitched notes throughout long segments of pieces give a stronger essence of a “horizontal” over a “vertical” harmony. Nonetheless, some primitive experimental results on automatic harmonization with HMMs indicate that the GCT types and extensions of most probable chord successions, as reflected in a transition matrix, encapsulate the potential of efficient voice leading, allowing the successions of vertical harmonic blocks to

be combined in such a way that an efficient voice leading algorithm would potentially interpret some basic horizontal characteristics. Although this argument is clearly supported by the examples presented in the experimental results section, a more elaborate examination is left for future work. Additionally, as a subject of future research, the determination of appropriate voice leading combinations will be realized by an algorithm that acts on the consecutive vertical harmonic blocks.

3.1 Intermediate anchor chords as boundary constraints

The chords that “connect” two successive fixed-boundary chord segments are defined by the aforementioned variation of HMM, the BCHMM. Throughout the development of the BCHMM, a nomenclature relative to the subject under discussion will be followed, i.e. the dataset will comprise musical pieces, the states will represent chords and the observations will describe melody notes. To this end, the set of possible states-chords will be denoted by \mathcal{S} , while the letters C and c will be used for denoting chords. The set of all possible observations-notes will be denoted as \mathcal{Y} , while Y and y will be denoting melody notes. Specifically, the capitalized letters will be used to denote statistical variables, while their instantiation variables will be denoted by lower case letters. For example, $P(C_i = c_i)$ denotes the probability that the chord in the i -th position is a c_i chord (where c_i is a specific chord, e.g. a [7, [0, 4, 7], [10]] chord in GCT form, which is a dominant seventh chord).

An initial set of music phrases is considered which will provide the system with the required statistical background, constituting the training set. Through this dataset the statistics that are induced concern three aspects:

1. The probability for each state (chord) to be a beginning chord. This distribution is computed by examining each beginning chord for each piece in the dataset and is denoted as $\pi(C_1 = c)$, $c \in \mathcal{S}$.
2. The probability for each state (chord) to be an ending chord. This distribution is computed by examining each ending chord for each piece in the dataset and is denoted as $\tau(C_T = c)$, $c \in \mathcal{S}$.
3. The probability that each state follows another state, denoted as $P(C_i = c_i | C_{i-1} = c_{i-1})$, $c_i, c_{i-1} \in \mathcal{S}$.
4. The probability of a chord being played over a melody note, denoted as $P(C_i = c_i | Y_i = y_i)$.

These probabilities are related during the computation of the *overall* probability that a certain chord sequence ($C_i = c_i$, $i = 1, 2, \dots, T$) is applied over an observed melody ($Y_i = y_i$, $i = 1, 2, \dots, T$). This overall probability is computed by

$$P(C_i = c_i | Y_i = y_i) = P_\pi P_\mu P_\tau, \quad (1)$$

where

$$P_\pi = \pi(C_1 = c_1) P(C_1 = c_1 | Y_1 = y_1), \quad (2)$$

$$P_\mu = \prod_{i=2}^T P(C_i = c_i | C_{i-1} = c_{i-1}) P(C_i = c_i | Y_i = y_i), \quad (3)$$

$$P_\tau = \tau(C_T = c_T) P(C_T = c_T | Y_T = y_T). \quad (4)$$

An optimal sequence of chords is one that maximizes the overall probability (in Equation 1)¹, by achieving an optimal path of states that yield a maximal combination for the probabilities in all the counterparts (P_π , P_μ and P_τ), typically through the Viterbi [16] algorithm. The probabilities in P_π promote some chords as better solutions to begin the path of chords: the ones that are more often used in the beginning of pieces in the dataset. Similarly, the probabilities in P_τ advance solutions that are more often met as concluding chords. Although the results reported in past works indicate that P_π and P_τ most probably create satisfactory results, these probabilities do not *guarantee* that the more often met beginning and ending chords will be utilized. A similar comment can be made about some strategies that have been proposed, which focus on constructing satisfactory cadences, by beginning from the end of the phrase to be harmonized and employing the Viterbi algorithm from “right-to-left”. Specifically, while the latter approaches have an increased bias towards the cadence part of the phrase, it is again not guaranteed that the cadence or the beginning chord of the phrase will be satisfactory.

Regarding the probabilistic scheme, the process for computing the probability value in Equation 1, incorporates the extraction of the statistical values for $\pi(C_1 = c_1)$ and $\tau(C_T = c_T)$, according to the number of occurrences of each chord as an initial or final chord respectively. For the BCHMM approach however, no statistics are considered for these boundary points, since they *certainly* (with probability 1) include the chords specified by a higher hierarchical level or by a human annotator. To be compatible with the terminology followed hitherto for the presentation of the HMM model, the latter comment can be expressed by modifying the Equations 2 and 4 so that they indicate the chords selected at temporary boundary points between successive checkpoints as certain, while eliminating the probabilities for any other chords to appear. Specifically, if the beginning and ending chords are selected to be α_1 and α_T respectively, the new probabilities that substitute the ones expressed by Equations 2 and 4 are the respective following:

$$P'_\pi = \begin{cases} 1, & \text{if } C_1 = \alpha_1 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$$P'_\tau = \begin{cases} 1, & \text{if } C_T = \alpha_T \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

¹In implementations of HMMs it is usually the negative log-likelihood that is being minimized, i.e. the logarithm of the expression in Equation 1, since the numbers that are yielded by consecutive multiplications of probabilities (quantities ≤ 0) are difficult to be compared by eye because of their small magnitude.

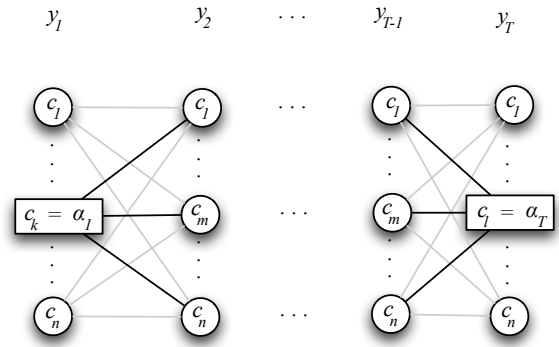


Figure 1. Trellis diagram for the BCHMM. Only transitions from α_1 and to α_T as first and last states respectively are permitted. The intermediate trellis diagram is the same as in a typical HMM.

The probability that is therefore optimized is the following:

$$P(C_i = c_i | Y_i = y_i) = P'_\pi P_\mu P'_\tau, \quad (7)$$

where the factor P_μ is the one defined in Equation 3. The employment of the Viterbi algorithm under the constraints imposed by the boundary conditions, as reflected by Equations 5 and 6, assigns zero-value probabilities to all paths, except the ones that begin with α_1 and end with α_T . Figure 1 illustrates the trellis diagram of the Viterbi algorithm under the discussed constraints.

3.2 Application of BCHMM in the current harmonization system

The efficiency of the HMM, and consequently the BCHMM, methodology relies on selecting a proper set of states to represent the chords that are utilized in the training set, which will subsequently be used in the harmonic generation process. The term “proper” indicates that there is a tradeoff in the amount of information of chord representation and the number of states required to delegate each chord in the HMM (and the BCHMM). For instance, by describing the possible chords only as major or minor, the number of states remains small (24 for all 12 pitch classes), however the harmonic description is very poor. Several works in the literature ([9, 15] among others) propose the utilization of standard chords (e.g major, minor, diminished, augmented and major seventh), applicable to all 12 relative pitch classes of the composition key of the examined pieces. However, by devising such a chord selection scheme it is possible that important harmonic information is excluded, since several pitch class combinations that might appear (rather frequently in some musical idioms) are disregarded.

The chord representation followed in the context of the paper at hand is the *general chord type* (GCT) representation, which is able to embody the information of both consonant and dissonant parts of a pitch class group. The GCT incorporates three parts, the *root*, the *type* and the *extensions* of a chord, denoted with three different entrances

in a list of the form [root, [type], [extension]]; for example the pitch class [7, 11, 2, 5] is represented as [7, [0, 4, 7], [10]], which indicates a dominant seventh chord. These parts are defined for pitch class simultaneities, according to a process that isolates the maximal mutually consonant pitch class combinations of this simultaneity, according to a consonance vector that defines the intervals between pitch classes that are considered consonant. For the chorales of Bach, that constitute the dataset of examination, the consonant intervals are considered to be the major and minor thirds, their inversion–equivalent major and minor sixths and the natural fifths and fourths. A complete description of the GCT is beyond the context of this paper and the interested reader is referred to [citation omitted for blind review].

The implementation of the HMM incorporated a simple “rule–based” observation–to–state probability assignment ($P(C_i = c_i | Y_i = y_i)$) for defining the probability for each chord to be played with each note of the melody. Specifically, for each note of the melody, this “rule–based” criterion provides a maximum probability for to chords that include this notes and a minimum for one that does not. To maximum probability is set to 1, while the minimum is set to 10^{-6} . Additionally, the zero entries of the chord transition matrices that are produced by the training simulations, are also assigned a value of 10^{-6} ². By removing the zero entries in these matrices, a potential blocking of the algorithm is avoided in situation where zero probabilities occur. Such situations may occur either in the extreme scenario where there is no chord to include a melodic note, or in the even more extreme scenario where there is no probable path connecting two predetermined anchor points.

4. RESULTS

The experimental results demonstrate in a qualitative manner the effectiveness concerning several aspects of the proposed melodic harmonization approach:

1. The effectiveness of the GCT representation towards capturing the idiom’s “chords”, providing interpretations that are in agreement with the Roman numeral analysis.
2. The efficient adaptation of the GCT to the chord types and extension characteristics that enable the automatic harmonization system to be amenable to effective voice leading. Dissonance of extensions, should be treated for special voice leading.
3. The presented methodology’s effectiveness in terms of the training data requirements.
4. The increase of interestingness that the insertion of intermediate and/or boundary chords can introduce to the composed harmony.

² After the adjustment of the values in either the observation or the transition matrices, these matrices could be normalized to produce a unit sum for each chord. However, since the probability values in the matrix entries are computed only in terms of maximizing the total probability (ignoring its magnitude), such a normalization is not necessary.

5. The fact that the HMMs are versatile enough to adapt to “deterministic” harmonic constraints.

During the “unofficial” evaluation of the presented methodology, several test phrases were harmonized, as well as several anchor point insertion setups were examined. The presented results include some indicative harmonizations that have been produced by the system with different anchor point setups. The utilized dataset comprises a selection of phrases from the “benchmark” chorales of J. S. Bach, specifically some chorales in the major mode.

The phrases of the Bach chorales were divided in two sets according to their key of composition, i.e. in major and minor phrases. Although harmonizations of both modes were tested, the reported results include only major mode phrases. The GCT chords–states that are derived for the major chorales of Bach are 41 and for the minor chorales 38, while many of the major and minor states are overlapping, i.e. exist both in the major and in the minor chorales. Several of these states are redundant since their GCT expression in fact describes chords of the same functionality, e.g. the GCTs [0, [0, 4, 7], []] and [0, [0, 4], []] denote a major chord in the tonic. Additionally, there is a considerable amount of GCT states (around 15 for each mode) that occur only two or three times in the entire dataset. The latter comments indicate that the employment of a GCT clustering technique could group some GCTs according to their harmonic functionality, further reducing the states to approximately 25 for each mode. However, such a grouping methodology is yet to be developed and it is part of ongoing research.

When harmonizing a melody with no constraints, the HMM methodology selects the most probable sequence of chords (hidden states) according to probabilities related to the melody’s note to be harmonized and to probabilities related to the transitions between pairs of states. The imposition of fixed–chord constraints is intuitively expected to alter the harmonization “locally”, i.e. the CHMM harmonization is expected to be different than the one provided by the typical HMM a few chords before or after a chord that remains fixed – if the selected chord to be fixed is different than the one provided by the HMM. However, the application of chord constraints in some cases provided different harmonizations throughout the entire length of the phrase. The voice leading in the examples presented below was performed by a music expert; an algorithmic process for voice leading is a future research goal.

The example in Figure 2 amplifies the role of anchor chords and specifically the beginning and ending chords of a phrase. In this example, a Bach chorale melody is harmonized with the typical HMM methodology (top) and with anchor boundary (beginning and ending) chords denoted by an asterisk. The boundary chords are the ones utilized by Bach in the genuine chorale. An initial comment concerns the fact that the HMM methodology does not “guarantee” that the beginning and ending (boundary) chords of a melody to be harmonized are identical to the ones that would potentially be utilized by a human composer. Additionally, the role of the boundary chords is crucial: the example Figure 2 demonstrates that different anchor chords

provided an entirely different harmonization. Furthermore, this example shows that the imposition of constraints “forced” the system to follow more “interesting” and unpredictable chord paths, since, the typical HMM methodology utilized more typical and probable chord progressions between V and I chords. The imposition of constraints on the other hand, forced the HMM methodology to establish temporary secondary tonalities, yielding a richer harmonic interpretation of the melodic sequence.

(a) typical HMM

(b) CHMM with boundary anchor chords

Figure 2. (a) The harmonization of a Bach chorale melody with the typical HMM methodology and (b) with constraints on the first and final chords (indicated with an asterisk).

The evidently important role of the beginning and ending chords leads to further inquiries about the ability of the HMM to accurately “predict” the boundary chords of phrases, according to the ones utilized in the genuine compositions. Answers to these inquiries are approached through a statistical comparison between the boundary chords produced by the HMMs and the boundary chords assigned by Bach. Specifically, an intuitively realistic answer is pursued with the utilization of three different metrics on how “correct” the boundary chords attributed by the HMM are, considering the boundary chords of the genuine Bach chorales phrases as ground-truth. Specifically, when the HMM system harmonizes the melody of a phrase, the attributed *first* and *final* GCT chords of the HMM harmonization are compared (according to the aforementioned three metrics) with the respective GCT chords that exist in the genuine harmonization of Bach on the same phrase. Therefore, these three metrics are considered to indicate the “efficiency” of the HMM harmonization regarding the beginning and ending GCT chords. These metrics are the following:

1. *Pitch class similarity* (PC, $\in [0, 1]$): the percentage of pitch classes (PCs) in the HMM proposed chord that are equal to the pitch classes of the “correct” chord.
2. *Root similarity* (root, $\in \{0, 1\}$): 1 if the GCT roots are equal, 0 otherwise.
3. *Exact similarity* (exact, $\in \{0, 1\}$): 1 if the GCT chords are completely equal, 0 otherwise.

The PC criterion is the most indulgent one, since it provides a rather positive score to chords that are considered wrong. For example, if the final chord in a phrase is $[0, 4, 7]$, $[\]$ and the HMM proposes an arguably wrong $[4, 0, 3, 7]$, $[\]$ chord, then it receives a score of 0.6667, since the common relative to the root PCs are 4 and 7, while the non-common is only the relative PC 11 (contradicting to 0). The exact criterion is the strictest criterion, since it requires that the root, type and extension between chords are the same. The root criterion admits that it is an exaggerative requirement that all the GCT chord characteristics be the same, acknowledging also the fact that potentially different GCT types and extensions refer to chords of the same functionality, e.g. $[0, [0, 4, 7], [\]]$ and $[0, [0, 4], [\]]$. To this end, the root criterion accounts only the similarity of the root GCT part.

The experimental setup includes four different sets of training excerpts, namely the $tr - 5$, $tr - 10$, $tr - 20$ and $tr - 30$ sets. Each of these sets comprises a number of training phrases that is indicated by the numerical part of the name, e.g. the $tr - 20$ describes an experimental simulation where 20 phrases are used as training data. Under any training scenario, 10 test melodies are harmonized, which belong to chorale phrases that do not pertain to the training set. The training and testing chorales are randomly selected in 100 random selection–training–harmonizing–testing simulations, while different sessions are performed for major and minor mode chorales. Thereby, the statistics that are subsequently presented are extracted from 100 simulations for each setup: major or minor chorale phrases, with different numbers of training phrases (5, 10, 20 and 30) and 10 phrases as harmonizing–testing data.

Table 2 demonstrates the mean values for the three efficiency measures in the first and final chords of the HMM harmonizations, for the major and the minor chorales and for all training setups (different number of training pieces). A first comment concerns the sensitivity of each metric to the number of training pieces. For instance, the PC metric remains relatively steady regardless of the number of pieces as a training set, while the remaining two metrics increase considerably as the number of training pieces increase. Specifically, for the major pieces the increase is around 10%, while for the minor piece around 4-5%. This fact indicates that the number of coinciding pitch classes is a rather vague measure, incorporating little musical information, since this measure does not reveal the dense impact that the increase of the training data would expectedly have.

Except from the imposition of boundary chords, the insertion of intermediate chords can also produce interesting results. The example depicted in Figure 3 discusses the harmonization of a Bach chorale in four different versions. Specifically, Figure 3 (a) demonstrates the harmonization produced by the typical HMM methodology, while the harmonization in (b) is produced with constraints on the boundary chords (as indicated by the asterisks). The constraints used in the phrase’s boundaries are the ones utilized by Bach in the genuine chorales. The imposition of the boundary constraints does not produce a harmonization that is

Table 2. Efficiency of the typical HMM harmonization regarding the first and final chords, according to the three defined metrics.

	Beginning		Ending	
	major	minor	major	minor
	pitch class similarity (PC)			
tr-5	0.8635	0.7917	0.9373	0.8777
tr-10	0.8698	0.8014	0.9437	0.8827
tr-20	0.8650	0.8008	0.9358	0.8934
tr-30	0.8670	0.7970	0.9533	0.8884
	root similarity (root)			
tr-5	0.4820	0.4110	0.6860	0.7330
tr-10	0.4940	0.4060	0.7460	0.7720
tr-20	0.4900	0.4420	0.7770	0.8010
tr-30	0.5310	0.4380	0.8230	0.7840
	exact matches (exact)			
tr-5	0.4360	0.3370	0.6530	0.4580
tr-10	0.4530	0.3280	0.7220	0.4990
tr-20	0.4660	0.3740	0.7570	0.5020
tr-30	0.5120	0.3710	0.7920	0.4980

entirely different regarding the selection of GCT chords (unlike the example shown in Figure 2), however the voice leading that was assigned by the music expert in both phrases is different. The harmonization became more interesting when the music expert indicated the insertion of the diminished chord marked with an asterisk in Figure 3 (c) (fifth chord). This anchor chord changed the harmonization entirely; even when the boundary constraints were alleviated, the harmonization produced by the CHMM system (Figure 3 (d)) was again completely novel. The fact that different constraint conditions produce diverse harmonizations, amplifies the motivation to utilize a “deterministic” chord selection scheme along with the probabilistic HMM framework.

5. CONCLUSIONS

The paper at hand presents a methodology for performing automatic melodic harmonization, i.e. providing chords on the notes of a given melody, through a methodology that is based on the hidden Markov model (HMMs), namely the constrained HMM (CHMM), which harnesses the capabilities of the HMMs to perform harmonizations with strictly specific requirements expressed through the employment of certain chords to harmonize certain notes of a melody. Such “anchor” chords would be selected either by an algorithmic (probably non-probabilistic) process functioning at a higher level of the harmonic hierarchy, or by a user. The utilization of specific chords imminently enhances the automatically produced harmonizations since the proper selection of some key-chords leads the system to interesting harmonic paths. For instance, the selection of the first and final (boundary) chords of a phrase, which chords strongly imply the tonal constitution, is a crucial part for generating harmonizations that provide strong reference to an intended musical idiom – fact that is also highlighted by several works in the automatic harmonization literature.

(a) typical HMM

(b) CHMM with boundary anchor chords

(c) CHMM with boundary and intermediate anchor chords

(d) CHMM with an intermediate anchor chord

Figure 3. (a) The harmonization of a Bach chorale melody with the typical HMM methodology and with constraints on (b) the boundary chords, (c) the boundary and one intermediate chord and (d) only one intermediate chord. The fixed intermediate chords selected by a human annotator are indicated on the score with an asterisk.

According to the experimental results reported in this paper, the typical HMM approach assigns beginning and ending chords of phrases that are more probable, a fact that potentially contradicts with a composer’s choices. Additionally, the imposition of fixed-chord constraints, even only on the boundaries of phrases, force the CHMMs to produce harmonizations that are completely different to the ones produced without constraints – an even more interesting since they are more “improbable”. The chord representation that is employed is the general chord type (GCT) representation, which is a novel technique under development and allows the selection of a relatively small number of chords as states, without disregarding harmonic information from chord extensions.

The proposed technique is a part of an ongoing research in the context of the COINVENT project, according to which the invention of new concepts in automated harmonization is approached by blending harmonic concepts of several musical idiom. To this end, the determination and utilization of important harmonic parts of idioms is pursued, e.g. selecting proper fixed-chord constraints (“anchor” chords) and voice leading among others. Therefore, the proposed technique remains to be integrated with an algorithmic “anchor” chord selection mechanism, as well as an algorithmic process that performs idiom-dependent

voice leading. The development of the CHMM methodology would potentially be harnessed with even more advanced and abstract harmonic constraints. For example, the user of a system would not only select entire chords to harmonize certain notes of phrases, but also specific notes that should be present along with a note of a harmony, therefore reducing the chord possibilities. Additionally, as the results indicated, by “fixing” the final boundary point it is not expected to lead to a “fixed” cadential pattern, since the absolute similarity in the final chord between the genuine and the artificial harmonies was not followed by an increase to the pre-final chords. The utilization of longer harmonic segments in places where cadences happen has been previously discussed in the literature [10], providing pointers for future work that would include larger cadential “chunks” as ending boundary points. Finally, the boundary constrained formalization could be harnessed with a variable order Markov model in the hidden layer, like the predictions suffix trees, producing results by potentially incorporating information over longer harmonic parts for deciding the next chords.

Acknowledgments

This work is funded by the COINVENT project. The project COINVENT acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 611553.

6. REFERENCES

- [1] K. Ebcioğlu, “An expert system for harmonizing four-part chorales,” *Computer Music Journal*, vol. 12, no. 3, pp. 43–51, 1988.
- [2] F. Pachet and P. Roy, “Musical harmonization with constraints: A survey,” *Constraints*, vol. 6, no. 1, pp. 7–19, Jan. 2001.
- [3] P. Donnelly and J. Sheppard, “Evolving four-part harmony using genetic algorithms,” in *Proceedings of the 2011 International Conference on Applications of Evolutionary Computation - Volume Part II*, ser. EvoApplications’11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 273–282.
- [4] S. Phon-amnuaisuk and G. A. Wiggins, “The four-part harmonisation problem: A comparison between genetic algorithms and a rule-based system,” in *Proceedings of the AISB99 symposium on musical creativity*. AISB, 1999, pp. 28–34.
- [5] S. Suzuki, “Four-part harmonization using probabilistic models: Comparison of models with and without chord nodes,” Stockholm, Sweden, pp. 628–633, 2013.
- [6] R. P. Whorley, G. A. Wiggins, C. Rhodes, and M. T. Pearce, “Multiple viewpoint systems: Time complexity and the construction of domains for complex musical viewpoints in the harmonization problem,” *Journal of New Music Research*, vol. 42, no. 3, pp. 237–266, Sep. 2013.
- [7] S. A. Raczyski, S. Fukayama, and E. Vincent, “Melody harmonization with interpolated probabilistic models,” *Journal of New Music Research*, vol. 42, no. 3, pp. 223–235, 2013.
- [8] J.-F. Paiement, D. Eck, and S. Bengio, “Probabilistic melodic harmonization,” in *Proceedings of the 19th International Conference on Advances in Artificial Intelligence: Canadian Society for Computational Studies of Intelligence*, ser. AI’06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 218–229.
- [9] C.-H. Chuan and E. Chew, “A hybrid system for automatic generation of style-specific accompaniment,” in *Proceedings of the 4th International Joint Workshop on Computational Creativity*. Goldsmiths, University of London, 2007.
- [10] N. Borrel-Jensen and A. Hjortgaard Danielsen, “Computer-assisted music composition – a database-backed algorithmic composition system,” B.S. Thesis, Department of Computer Science, University of Copenhagen, Copenhagen, Denmark, 2010.
- [11] M. Allan and C. K. I. Williams, “Harmonising chorales by probabilistic inference,” in *Advances in Neural Information Processing Systems 17*. MIT Press, 2004, pp. 25–32.
- [12] M. Hanlon and T. Ledlie, “Cpu bach: An automatic chorale harmonization system,” 2002. [Online]. Available: <http://www.timledlie.org/cs/CPUBach.pdf>
- [13] L. Yi and J. Goldsmith, “Automatic generation of four-part harmony,” in *BMA*, ser. CEUR Workshop Proceedings, K. B. Laskey, S. M. Mahoney, and J. Goldsmith, Eds., vol. 268. CEUR-WS.org, 2007.
- [14] N. Yogev and A. Lerch, “A system for automatic audio harmonization,” 2008.
- [15] I. Simon, D. Morris, and S. Basu, “Mysong: Automatic accompaniment generation for vocal melodies,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’08. New York, NY, USA: ACM, 2008, pp. 725–734.
- [16] J. Forney, G.D., “The viterbi algorithm,” *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.