Learning and blending harmonies in the context of a melodic harmonisation assistant

Maximos Kaliakatsos-Papakostas, Dimos Makris*, Asterios Zacharakis, Costas Tsougras, and Emilios Cambouropoulos

Department of Music Studies, Aristotle University of Thessaloniki,
54124, Thessaloniki, Greece
*Department of Informatics Studies, Ionian University,
49100, Corfu, Greece
{maxk,aszachar,tsougras,emilios}@mus.auth.gr
c12makr@ionio.gr

Abstract. How can harmony in diverse idioms be represented in a machine learning system and how can learned harmonic descriptions of two musical idioms be blended to create new ones? This paper presents a creative melodic harmonisation assistant that employs statistical learning to learn harmonies from human annotated data in practically any style, blends the harmonies of two user-selected idioms and harmonises user-input melodies. To this end, the category theory algorithmic framework for conceptual blending is utilised for blending chord transition of the input idioms, to generate an extended harmonic idiom that incorporates a creative combination of the two input ones with additional harmonic material. The results indicate that by learning from the annotated data, the presented harmoniser is able to express the harmonic character of diverse idioms in a creative manner, while the blended harmonies extrapolate the two input idioms, creating novel harmonic concepts.

1 Introduction

Machine learning allows a machine to acquire knowledge from data forming concrete conceptual spaces, while conceptual blending [10] between two input spaces allows new spaces to be constructed expressed as new structural relations or even new elements, creating new and potentially unforeseen output [27]. In music, harmony is an characteristic and well-circumscribed element of an idiom that can be learned from human annotated musical data using techniques such as Hidden Markov Models, N-grams, probabilistic grammars, inductive logic programming (see [25, 22, 6, 23, 20, 21, 26, 7, 12, 17, 15] among others). In the context of computational creativity in music, a challenging task tackled in the Concept Invention Theory (COINVENT) [24, 18, 3] project is to blend different/diverse input harmonic idioms learned from data to create new idioms that are creative supersets of the input ones.

The paper at hand briefly presents the extension of a melodic harmonisation assistant (introduced in [15]) that learns harmonic idioms by statistical learning

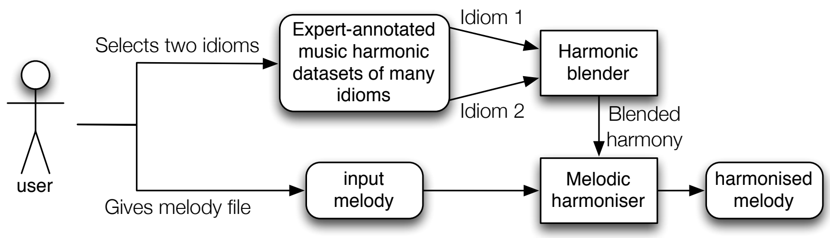


Fig. 1. System overview.

on human data, for inventing new harmonic spaces by blending the 'atoms' of harmony, i.e. transitions between chords. The blended transitions are created by combining the features characterising pairs of transitions belonging to two idioms (expressed as sets of potentially learned transitions) according to an amalgam-based algorithm [9, 5] that implements the theory presented in [10] for conceptual blending, through the categorical-based methodology presented in [11]. The transitions are then used in an extended harmonic space that accommodates the two initial harmonic spaces, linked with the new blended transitions.

2 Statistical learning of harmonies from human annotated datasets

Before blending harmonies, the system learns different aspects of harmony through annotated training data, while it produces new harmonisations according to guidelines provided in a melody input file given by the user. The system learns the available chord types within diverse dataset (according to their root notes) based on the *General Chord Type* (GCT) [2] representation, which can be used not only to represent but also to describe meaningful relations between harmonic labels [16] – even in non-tonal music idioms [14, 1]. The training data include musical scores from many idioms, with expert annotations. Specifically, the notes of harmonic manually annotated reductions are regarded for the harmonic learning process, where only the most important harmonic notes are included; additional annotated layers of information are given regarding the tonality and the metric structure of each piece. Accordingly, the format of the user melody input file includes indications of several desired attributes that the resulting harmonisation should have.

By employing rich multi-level structural descriptions of harmony in different idioms, as included in 'real-world' representations of historical traditions, the harmoniser is able to create new music that accurately reflects the characteristics of these idioms. A diverse collection of musical pieces drawn from different historic eras and from different harmonic styles has been assembled by music experts. The dataset consists of over 430 manually annotated musicXML documents categorised in 7 categories, internally as coherent as possible, and various subcategories, presented in the following list:

1. Modal harmonisation in the Middle Ages (11th – 14th centuries): includes subcategories of medieval pieces in the Organum and Fauxbourdon styles.
2. Modal harmonisation in the Renaissance (15th – 17th centuries): includes modal music from the 16th – 17th centuries along with modal chorales.
3. Tonal harmonisation (17th – 19th centuries): includes a set of the Bach Chorales, the Kostka-Payne corpus and tonal harmonisation sets from the 18th – 19th centuries.
4. Harmonisation in National Schools (19th – 20th centuries): includes 19th – 20th century harmonisation of folk songs from Norway, Hungary and Greece.
5. Harmonisation in the 20th century: includes harmonisations of Debussy, Hindemith, Whitacre, Stravinsky and Schnittke among others.
6. Harmonisation in folk traditions: includes Tango (classical and nuevo styles), Epirus polyphonic songs and Rebetiko songs.
7. Harmonisation in 20th-century popular music and jazz: includes mainstream jazz, pieces from Bill Evans and a collections of songs from The Beatles.

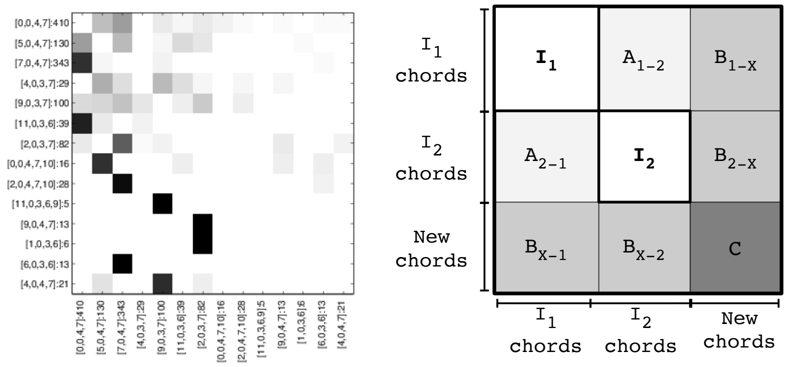
The musical pieces are manually annotated by music experts such that structural harmonic features may be extracted at various hierarchic levels. Annotated music files include the original musical data the actual musical surface, along with manually entered expert annotations. *Expert annotations* refer to specific analytic concepts, e.g., the use of harmonic symbols to describe note simultaneities, modulations, phrasing etc. Specifically, the expert annotations are given in musical form and include time-span reduction of the original content (ms_1), annotations concerning *tonality* (and tonality changes) and *grouping* information. On the chord transitions level the system is trained according to the GCT-form chord progressions on the harmonic reduction (ms_1); the GCT forms of the extracted chords are computed by using the tonality annotations. Annotations of grouping indicate the positions of phrase endings, where the system learns cadences (the final pairs of chords).

After the system is trained, it is able to harmonise a given melody given, in this stage, a manually annotated input melodic file. The melodic annotations concern harmonic rhythm, harmonically important notes, key and phrase structure. The user provides the harmonic rhythm as the positions where chords should occur and the important notes (harmonic notes) that should be considered with higher weight when selecting chords for each segment. If the user provides no information for these attributes, the system produces default harmonic rhythm and important note selection schemes that might lead to 'unwanted' harmonic results. Additionally, the user has the freedom to choose specific chords at desired locations (constraint chords), forcing the system creatively to produce chord sequences that comply with the user-provided constraints, therefore allowing the user to 'manually' increase the interestingness of the produced output. Finally, the user should accompany the melody with higher level harmonic information concerning the tonality or tonalities of the piece, as well as with its phrase grouping boundaries. Tonality is indicated by a cluster of all notes included in the scale, with the lowest note indicating the tonality's tonic. Grouping is annotated by arbitrary notes at the metric position where grouping changes occur, while the number of notes in these positions indicate the grouping level of the phrase.

The cHMM [17] algorithm is used for modelling/learning chord progression probabilities for a given idiom. Then statistical information from the user-defined melody is combined with the chord progression model to generate chord progressions that best represent the idiom. Additionally the algorithm offers the possibility for prior determination of intermediate 'checkpoint' chords [4]). The fixed intermediate chords on the one hand help towards preserving some essence of higher level harmonic structure through the imposition of intermediate and final cadences, while on the other hand allow interactivity by enabling the user to place desired chord at any position. Statistics for cadences are learned during the training process, where expert annotated files including annotations for phrase endings are given as training material to the system. After collecting the statistics about cadences from all idioms, the system, before employing the cHMM algorithm, assigns cadences as fixed chords to the locations indicated by user input. The cadence to be imported is selected based on three criteria: (a) whether it is a final or an intermediate cadence; (b) the cadence likelihood (how often it occurs in the training pieces); and (c) how well it fits with the melody notes that are harmonised by the cadence's chords. Direct human intervention allows the user of the system to specify a harmonic 'spinal chord' of anchor chords that are afterwards connected by chord sequences that give stylistic reference to a learned idiom.

Regarding voice leading, experimental evaluation of methodologies that utilise statistical machine learning techniques demonstrated that an efficient way to harmonise a melody is to add the bass line first [26]. The presented harmoniser, having defined the optimal sequence of GCT chords, uses a modular methodology for determining the bass voice leading presented in [19], which utilises independently trained modules that include (a) a hidden Markov model (HMM) deciding for the bass contour (hidden states), given the melody contour (observations); (b) distributions on the distance between the bass and the melody voice; and (c) statistics regarding the inversions of the chords in the given chord sequence.

The bass voice motion provides abstract information about the motion of the bass, however, assigning actual pitches for a given set of GCT chords requires additional information: *inversions* and *melody-to-bass* distance distributions are also learned from data. The inversions of a chord play an important role in determining how eligible is each chord's pitch class to be a bass note, while the melody-to-bass distance captures statistics about the pitch height region that the bass voice is allowed to move according to the melody. After obtaining the exact bass pitch, the exact voicing layout, i.e. exact pitches for all chord notes, for each GCT chord is defined. To this end, a simple statistical model is utilised that finds the best combination of the intermediate voices for every chord according



(a) Major Bach chorales (b) Extended matrix

Fig. 2. Graphical description of (a) the transition matrix in a set of major-mode Bach chorales and (b) an *extended* matrix that includes transition probabilities of two initial idioms – like the ones depicted in (a) – and of several new transitions generated through transition blending.

to some simple criteria. These criteria include proximity to a *pitch-attractor*, evenness of *neighbouring notes distances* and inner voice *movement distances between chords*. These criteria form an aggregate wighted sum that defines the optimal setting for all the intermediate notes (between the bass and the melody) in every GCT chord.

3 Blending learned harmonies

In the presented system, the harmony of an idiom is represented by first order Markov matrices, which include one respective row and column for each chord in the examined idiom. The probability value in the i-th row and the j-th column exhibits the probability of the i-th chord going to the j-th —the probabilities of each row sum to unit. Figure 2 (a) illustrates a grayscale interpretation of the transition in a set of major-mode Bach chorales. An important question is: *Given two input idioms as chord transition matrices, how would a blended idiom be expressed in terms of a transition matrix?* The idea examined in the present system is to create an *extended* transition matrix that includes new transitions that allow moving across chords of the initial idioms by potentially using new chords. The examined methodology uses *transition blending* to create new transitions that incorporate blended characteristics for creating a smooth 'morphing' harmonic effect when moving from chords of one space to chords of the other. An abstract illustration of an extended matrix is given in Figure 2 (b).

Transition matrix is performed through *amalgam-based* conceptual blending that has already been applied to invent chord cadences [8, 28]; in this setting, cadences are considered as special cases of chord transitions—pairs of chords, where the first chord is followed by the second one— that are described by means

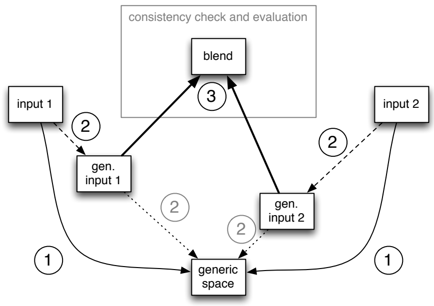


Fig. 3. Conceptual blending based on amalgam. The generic space is computed (1) and the input spaces are successively generalised (2), while new blends are constantly created (3). Some blends might be inconsistent or purely evaluated according to blending optimality principles or domain specific criteria.

of features such as the roots or types of the involved chords, or intervals between voice motions, among others. When blending two transitions, the amalgam-based algorithm first finds a generic space between them (point 1 in Figure 3), which is the set of their common features. Afterwards, the amalgam-based process computes the amalgam of two input spaces by *unifying* their content. If the resulting amalgam is inconsistent, then it iteratively generalises the properties of the inputs (point 2 in Figure 3), until the resulting unification is consistent (point 3 in Figure 3). Therefore, the amalgam-based process generalises the clashing property in one of the inputs and tries to unify the generalised versions of the inputs again. After a number of generalisation steps are applied (point 2 in Figure 3), the resulting blend is consistent (point 3 in Figure 3). For more information on transition blending the reader is referred to [13].

By analysing the graphical representation of an extended matrix as depicted in Figure 2 the following facts are highlighted:

1. By using transitions in $\mathtt{I_i}$, only chords of the i-th idiom are used. When using these transitions, the resulting harmonisations preserve the character of idiom i.
2. Transitions in $\mathtt{A_{i-j}}$ create direct jumps from chords of the i-th to chords of the j-th idiom. Blended transitions in $\mathtt{A_{i-j}}$ can be directly included in the extended matrix.
3. Transitions in $\mathtt{B_{i-X}}$ constitute harmonic motions from a chord of idiom i to a newly created chord by blending. Similarly, transitions in the $\mathtt{B_{X-j}}$ arrive at chords in idiom j from new chords. For moving from idiom i to idiom j using one external chord c_x that was produced by blending, a chain of two transitions is needed: $c_i \rightarrow c_x$ followed by a transition $c_x \rightarrow c_j$, where c_i in idiom i and c_j in idiom j respectively. A chain of two consecutive transitions with one intermediate external chord from chords of i to chords of j will be denoted as $\mathtt{B_{i-X-j}}$.

4. Sector `C` transitions are disregarded since they incorporate pairs of chords that exist outside the i-th and j-th idioms, violating our hypothesis for moving from one known chord sets to the other using one new chord at most.

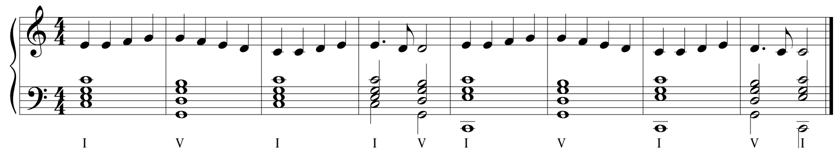
Based on this analysis of the extended matrix, a methodology is proposed for using blends between transitions in $\mathtt{I}_1$ and $\mathtt{I}_2$. Thereby, transitions in $\mathtt{I}_1$ are blended with ones in $\mathtt{I}_2$ and a number of the best blends is stored for further investigation, creating a pool of best blends. Based on multiple simulations, a large number of the best blends (i.e. 100) in each blending simulation should be inserted in the pool of best blends, so that several commuting scenarios can be created between the initial transition spaces. Thus, a greater number of blends in the pool of best blends introduces a larger number of possible commuting paths in $\mathtt{A}_{\mathtt{i-j}}$ or in $\mathtt{B}_{\mathtt{i-X-j}}$.

4 Harmonisation examples

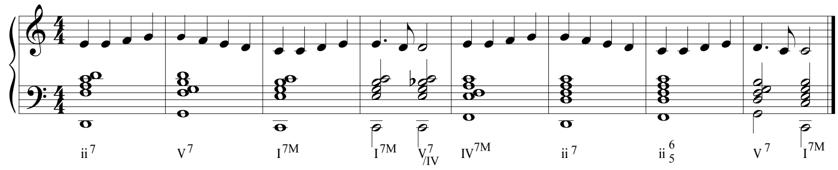
To briefly demonstrate the effect that transition blending has on forming the extended matrix that combines two initial idioms, harmonisations of the first part of '*Ode to Joy*' by Beethoven are illustrated in Figure 4. Initially, one can observe that the learned harmonic features from the Bach chorales and the jazz sets (Figure 4 (a) and (b) respectively) are reflected in the harmonisations that the system produces. In the case of the Bach chorales, the most convenient (yet not so musically impressive) sequence of chords is generated, where the V-I pattern is repeated. The jazz harmonisation includes modifications of the usual ii-V-I pattern. By blending the transitions of the two initial idioms, the produced harmonisation (illustrated in Figure 4 (c)) features new structural relations, incorporating chords and chord sequences that are not usual in any of the initial idioms. However, even thought the chords and chord transitions *per se* are unusual, their encompassed features reflect musical attributes of the initial idioms.

5 Conclusions

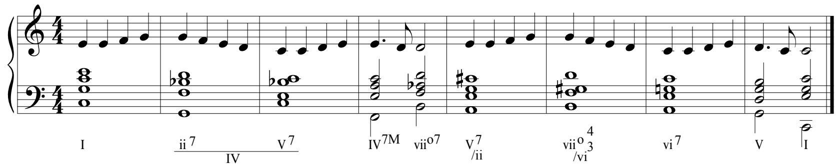
This paper describes a melodic harmonisation system which receives as inputs a melody file and two harmonic idioms and produces a melodic harmonisation with the blended harmony of two input idioms. To this end, diverse harmonic datasets were compiled and annotated by experts, while the harmonic description of each idiom was based on chords extracted with the General Chord Type (GCT) algorithm, statistical learning of chord progressions, cadences and chord voicing over these data. Blending of harmonies is performed through blending chord transitions (one chord leading to another) from the input idioms using an algorithmic realisation of conceptual blending based on category theory. Chord transition blending combines features between pairs of transitions from the two input harmonic idioms, producing new transitions that potentially include new



(a) *Ode to joy* harmonised in the style of Bach chorales.



(b) *Ode to joy* harmonised in the style of jazz.



(c) *Ode to joy* harmonised in the the blended style of Bach chorales and jazz.

Fig. 4. Beethoven's *Ode to joy* theme harmonised by the system with learned idioms and their blend.

chords for both idioms and incorporate blended features. These new blended transitions act as connection points between the two input harmonic idioms, generating the *extended idiom* that is a blended harmonic superset of the two input ones.

A thorough experimental process that evaluates the usefulness of the produced harmonisations in real-world applications (e.g. when the system is used as an assistant for composers) is underway. Initial experimental results indicate that the blended melodic harmonisations are more interesting than the ones produced by using each input idiom separately. Additional experimental processes are expected to provide insights into whether the blended harmonic spaces are perceived as alterations of one of the input spaces (one-sided blends), balanced blends or radically new harmonic idioms, as well as into the role of the user-defined melody in the process of using blended or non-blended harmonising idioms.

Acknowledgments

This work is funded by the Concept Invention Theory (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.

References

1. Cambouropoulos, E.: The Harmonic Musical Surface and Two Novel Chord Representation Schemes. In: Meredith, D. (ed.) Computational Music Analysis, pp. 31–56. Springer (2015)
2. Cambouropoulos, E., Kaliakatsos-Papakostas, M., Tsougras, C.: An idiom-independent representation of chords for computational music analysis and generation. In: Proceeding of the joint 11th Sound and Music Computing Conference (SMC) and 40th International Computer Music Conference (ICMC). ICMC–SMC 2014 (2014)
3. Cambouropoulos, E., Kaliakatsos-Papakostas, M., Tsougras, C.: Structural blending of harmonic spaces: a computational approach. In: Proceedings of the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM) (2015)
4. Chuan, C.H., Chew, E.: 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)
5. Confalonier, R., Schorlemmer, M., Plaza, E., Eppe, M., Kutz, O., Peñaloza, R.: Upward Refinement for Conceptual Blending in Description Logic – An ASP-based Approach and Case Study in $\mathcal{EL}^{++}$. In: International Workshop on Ontologies and Logic Programming for Query Answering, International Joint Conference on Artificial Intelligence (IJCAI) 2015 (2015)
6. Conklin, D.: Representation and discovery of vertical patterns in music. In: Music and artificial intelligence, pp. 32–42. Springer (2002)
7. Dixon, S., Mauch, M., Anglade, A.: Probabilistic and logic-based modelling of harmony. In: Exploring Music Contents, pp. 1–19. Springer (2010)
8. Eppe, M., Confalonier, R., Maclean, E., Kaliakatsos-Papakostas, M., Cambouropoulos, E., Schorlemmer, M., Codescu, M., Kühnberger, K.U.: Computational invention of cadences and chord progressions by conceptual chord-blending. In: International Joint Conference on Artificial Intelligence (IJCAI) 2015. p. Submitted (2015)
9. Eppe, M., Maclean, E., Confalonieri, R., Kutz, O., Schorlemmer, M., Plaza, E.: Asp, amalgamation, and the conceptual blending workflow. In: Calimeri, F., Ianni, G., Truszczynski, M. (eds.) Logic Programming and Nonmonotonic Reasoning, Lecture Notes in Computer Science, vol. 9345, pp. 309–316. Springer International Publishing (2015)
10. Fauconnier, G., Turner, M.: The Way We Think: Conceptual Blending And The Mind's Hidden Complexities. Basic Books, New York, reprint edition edn. (Mar 2003)
11. Goguen, J.: Mathematical Models of Cognitive Space and Time. In: Andler, D., Ogawa, Y., Okada, M., Watanabe, S. (eds.) Reasoning and Cognition, Interdisciplinary Conference Series on Reasoning Studies, vol. 2. Keio University Press (2006)
12. Granroth-Wilding, M., Steedman, M.: A robust parser-interpreter for jazz chord sequences. Journal of New Music Research 0(0), 1–20 (2014)
13. Kaliakatsos-Papakostas, M., Confalonier, R., Corneli, J., Zacharakis, A., Cambouropoulos, E.: An argument-based creative assistant for harmonic blending. In: Proceedings of the 7th International Conference on Computational Creativity (ICCC). p. (accepted) (2016)

14. Kaliakatsos-Papakostas, M., Katsiavalos, A., Tsougras, C., Cambouropoulos, E.: Harmony in the polyphonic songs of epirus: Representation, statistical analysis and generation. In: 4th International Workshop on Folk Music Analysis (FMA) 2014 (June 2014)
15. Kaliakatsos-Papakostas, M., Makris, D., Tsougras, C., Cambouropoulos, E.: Learning and creating novel harmonies in diverse musical idioms: An adaptive modular melodic harmonisation system. Journal of Creative Music Systems 1(1), (In press) (2016)
16. Kaliakatsos-Papakostas, M., Zacharakis, A., Tsougras, C., Cambouropoulos, E.: Evaluating the General Chord Type representation in tonal music and organising GCT chord labels in functional chord categories. In: Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2015). Malaga, Spain (2015)
17. Kaliakatsos-Papakostas, M., Cambouropoulos, E.: Probabilistic harmonisation with fixed intermediate chord constraints. In: Proceeding of the joint 11th Sound and Music Computing Conference (SMC) and 40th International Computer Music Conference (ICMC). ICMC–SMC 2014 (2014)
18. Kaliakatsos-Papakostas, M., Cambouropoulos, E., Kühnberger, K.U., Kutz, O., Smaill, A.: Concept Invention and Music: Creating Novel Harmonies via Conceptual Blending. In: In Proceedings of the 9th Conference on Interdisciplinary Musicology (CIM2014). CIM2014 (December 2014)
19. Makris, D., Kaliakatsos-Papakostas, M., Cambouropoulos, E.: Probabilistic modular bass voice leading in melodic harmonisation. In: Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2015). Malaga, Spain (2015)
20. Pérez-Sancho, C., Rizo, D., Inesta, J.M.: Genre classification using chords and stochastic language models. Connection science 21(2-3), 145–159 (2009)
21. Raphael, C., Stoddard, J.: Functional harmonic analysis using probabilistic models. Computer Music Journal 28(3), 45–52 (2004)
22. Rohrmeier, M.: Towards a generative syntax of tonal harmony. Journal of Mathematics and Music 5(1), 35–53 (Mar 2011)
23. Scholz, R., Vincent, E., Bimbot, F.: Robust modeling of musical chord sequences using probabilistic n-grams. In: Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. pp. 53–56. IEEE (2009)
24. Schorlemmer, M., Smaill, A., Kühnberger, K.U., Kutz, O., Colton, S., Cambouropoulos, E., Pease, A.: Coinvent: Towards a computational concept invention theory. In: 5th International Conference on Computational Creativity (ICCC) 2014 (June 2014)
25. Steedman, M.: The blues and the abstract truth: Music and mental models. Mental models in cognitive science pp. 305–318 (1996)
26. Whorley, R.P., Wiggins, G.A., Rhodes, C., Pearce, M.T.: Multiple viewpoint systems: Time complexity and the construction of domains for complex musical viewpoints in the harmonisation problem. Journal of New Music Research 42(3), 237–266 (Sep 2013)
27. Wiggins, G.A., Pearce, M.T., Müllensiefen, D.: Computational modeling of music cognition and musical creativity. In: The Oxford Handbook of Computer Music. OUP USA (2009)
28. Zacharakis, A., Kaliakatsos-Papakostas, M., Cambouropoulos, E.: Conceptual blending in music cadences: A formal model and subjective evaluation. In: ISMIR. Malaga (2015)