D7.3
Concept invention in melodic harmonisation and evaluation of melodic harmonisation assistant

Authors	Maximos Kaliakatsos-Papakostas, Asterios Zacharakis, Emilios Cambouropoulos
Reviewers	Allan Smaill, Ewen Maclean, Kai-Uwe Kuhnberger

Grant agreement no.	611553
Project acronym	COINVENT - Concept Invention Theory
Date	October 24, 2016
Distribution	PU

Disclaimer

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.

Abstract

This deliverable describes the research work for completing the Task 3 and Task 4 in the work package 7 of the COINVENT project. Task 3 involved the development of a melodic harmonisation assistant that learns harmonies of diverse idioms from expert-annotated data and constructs new harmonic spaces through blending. Task 4 focused on evaluating the products of this harmonisation assistant through subjective tests. The attached papers published in the context of this deliverable examined a wide range of harmonic blending aspects, including theoretic studies on blending in music, algorithmic approaches on blending and generating harmonies and experimental techniques for evaluating harmonic blending.
Keyword list: **Harmonic blending**, **Conceptual blending**, **Melodic Harmonisation**, **Harmonic Learning**, **Harmonic Blending Evaluation**

1 Executive Summary

Deliverable 7.3 incorporates creative approaches to harmonic generation through conceptual blending, focusing on the computational methodology created in the COINVENT project. With regards to music, conceptual blending has been predominantly theorised as the cross-domain integration of musical and extra-musical domains such as text or image and primarily discussed from a musico-analytical perspective focusing on structural and semantic integration. 'Intra-musical' conceptual blending in music is often conflated with the notion of structural blending and Fauconnier and Turner's theory is primarily applied to the integration of different or conflicting structural elements, such as chords, harmonic spaces, or even melodic-harmonic material from different idioms. The first task of this deliverable was to construct a melodic harmonisation assistant that generates harmonies that feature blended characteristics, using the COINVENT core model of conceptual blending; the second task was to evaluate its products.

Under this deliverable, several aspects of creative music blending were investigated in order to provide an extensive understanding of the tasks to be tackled. These aspects are divided in the following four research parts:

1. *Theoretic considerations of conceptual blending in music*: this part incorporated analytical research under the scope of conceptual blending in existing musical material, on the social dimension of creativity and further theoretical considerations on the nature of blending in music.

2. *Harmonic learning from data*: following-up the work of Deliverable 7.2, new learning components were proposed in order to enable statistical learning on several levels of harmony.

3. *Harmonic blending in a melodic harmonisation assistant*: in this part, the core model of conceptual blending, as developed in the COINVENT project, was employed for the automated generation of blends. Initial applications included chord blending, with interesting findings regarding cadences, while afterwards transition blending was integrated in a melodic harmonisation assistant that is capable of blending harmonic styles.

4. *Evaluating blending in harmony*: the interesting cadence results obtained by chord blending were further examined under an empirical setting, allowing perceptual evaluation of the blending results. Empirical evaluation on the products of the harmonisation assistant indicated that the intended purposes of blending were met, since the system produced perceivable blends that were equally preferred, compared to non-blends.

An additional outcome of the work under this deliverable is a forthcoming (3 September 2017) special issue in the 'Musicae Scientiae' journal, entitled 'Creative Conceptual Blending in Music', edited by Emilios Cambouropoulos, Danae Stefanou, Costas Tsougras (School of Music Studies, Aristotle University of Thessaloniki). This special issue will include contributions that cover a wide range of approaches on the utilisation of conceptual blending in music, expanding the discussion of generative conceptual blending and increasing the overall impact of the COINVENT project to the field of musical creativity. For matters of space, some representative publications related to the four aforementioned research parts are attached in this deliverable. For the full texts of all the publications related to this deliverable, as given in the list of references in this text, the interested reader is referred to the COINVENT website[1].

[1] `http://www.coinvent-project.eu/en/publicationsmedia/publications_by_type.html`

The theory of conceptual blending is based on a solid cognitive basis. Therefore, the applications and the extensions that conceptual blending has in music, were first studied under a theoretical point of view. In [14] a structural and hermeneutical analysis of 'Il vecchio castello' from Modest Musorgsky's 'Pictures at an Exhibition' in an attempt to disclose both the intra-musical (combination of modal, tonal and coloristic harmonic spaces) and the extra-musical (contextual, symbolic and programmatic aspects) conceptual blending that the work incorporates. The proposed analysis shows how musical structure promotes meaning construction through cross-domain mapping. This research suggests that conceptual blending theory as an analytical tool can promote a richer structural interpretation and experience of Musorgsky's work. The social aspect of social creativity, which is a crucial part of the COINVENT project, was examined in [12], where the theoretical and methodological developments in the study of social creativity in music were outlined, focusing on collaborative and improvised music-making. Particular reference was made to FolioHarmonies, a short qualitative study carried out as part of COINVENT project, and documenting collaborative, open-ended problem-solving processes in the creation of original musical pieces. Finally, a critical investigation of the application of Fauconnier and Turner's conceptual blending theory in music was presented in [13]. This study aimed to expose a series of questions and aporias highlighted by current and recent theoretical work in the field, related to the common distinction between intra- and extra-musical blending as well as the usually retrospective and explicative application of conceptual blending. It was thereby argued that more emphasis could be given to bottom-up, contextual, creative and collaborative perspectives of conceptual blending in music.

The creative melodic harmonisation assistant developed in the context of this deliverable, is able to learn harmonies from musical data of diverse harmonic idioms. Before presenting the blending methodology followed for harmonic blending and building on the developments presented in the Deliverable 7.2 of the COINVENT project, the methodologies used for learning harmonies were further examined and refined in the context of this deliverable. In [1] the General Chord Type (GCT) representation, which is the idiom-independent representation used by the creative melodic harmonisation assistant to learn and compose harmonies, was further investigated according to musicological, perceptual and computational aspects of the harmonic musical surface. The first step towards transforming abstract GCTs to actual notes incorporates an idiom-independent methodology for learning and composing the bass voice leading. In [10], a probabilistic approach was proposed for the automatic generation of voice leading for the bass note on a set of given chords from different musical idioms, according to a given set of GCT chord sequences. This methodology was based on the hidden Markov model (HMM) and the bass voice contour was determined by observing the contour of the melodic line. This approach was further refined in [11], by introducing additional statistical parameters for the motion of the bass voice except from the aforementioned HMM, namely: bass-to-melody distances as well as statistics regarding inversions and note doublings in chords. These parameters were incorporated in a modular approach, while the experimental results presented therein indicated that the proposed methodology captures rather effectively the voice layout characteristics of diverse idioms, whilst additional analyses on separate statistically trained modules revealed distinctive aspects of each idiom. The complete system for learning different harmonic aspects in diverse music idioms was presented in [5]. This system is adaptive (learns from data), general (can cope with any tonal or non-tonal harmonic idiom) and modular (learns different aspects of harmonic structure such as chord types, chord transitions, cadences and voice-leading) and can be used, not only to mimic given harmonic styles, but, to generate novel harmonisations for diverse melodies by exploring the harmonic possibilities provided by the implied harmonies of input melodies, or by allowing the imposition of user-defined chord constraints leading thus to new unforeseen harmonic realisations.

The creative applications of conceptual blending on many levels of harmony, were examined in [2] under the scope of structural, intra-musical blending. Specifically, the following aspects of structure were examined: chord-level blending, chord sequence blending, scale blending, harmonic structure level blending, melody-harmony level blending. Focussing on chord-level blending, the work presented in [3] used the COINVENT core model of conceptual blending based on amalgams to automatically find transitions between chord progressions of different keys or idioms and substitute chords in a chord progression by other chords of a similar function, as a means to create novel variations. The approach was the first evidence that demonstrated interesting creative blending examples, where jazz cadences are invented by blending chords in cadences from earlier idioms, and where novel chord progressions are generated by inventing transition chords. The chord blending paradigm was expanded to transition blending, which involved the introduction of many harmonic properties for describing transition ontologies. Deciding about the importance of such properties in the input spaces and evaluating the results of conceptual blending is a nontrivial task, specifically in the case of musical harmony where deep musicological background knowledge is required. In [4] a system was presented that allows a music expert to specify arguments over given transition properties, in a process that makes the system capable of defining combinations of features in an idiom-blending process. A music expert can assess whether the new harmonic idiom makes musicological sense and re-adjust the arguments (selection of features) to explore alternative blends that can potentially produce better harmonic spaces. The refined blending methodology that was developed in [4], was combined with the learning methodology presented in the previous paragraph and expanded in blending Markov transition matrices in the system presented in [6] and [7]. The melodic harmonisation assistant presented therein, features creative conceptual blending between two initial harmonic idioms, enabling various interesting music applications, ranging from problem solving, e.g. harmonising melodies that include key transpositions, to generative harmonic exploration, e.g. combining major-minor harmonic progressions or more extreme idiosyncratic harmonies.

The first step of evaluating the creative harmonic blending systems included the evaluation of the way it represents chords, namely through the GCT. In [8] the descriptive potential of the GCT was assessed in the tonal idiom by comparing GCT harmonic labels with human expert annotations (in the Kostka & Payne harmonic dataset) and the results indicated that the GCT representation constitutes a suitable scheme for representing effectively harmony in computational systems. Based on the interesting cadences that came out of the chord blending system, its ability to make fair predictions of the human-perceived dissimilarities between the blended cadences it produces was evaluated in [15]. This work was further expanded in [9], where the behavioural data were used as a 'ground-truth' human-based perceptual space of cadences, allowing an evolutionary algorithm to adjust the salience of each cadence feature to providing a system-perceived space of cadences that optimally matched the ground-truth space. This work was further expanded in [16], where a verbal attribute magnitude estimation method on six descriptive axes (preference, originality, tension, closure, expectancy and fit) is used to associate the dimensions of this space with descriptive qualities (closure and tension emerged as most prominent qualities). The novel cadences generated by the computational blending system were mainly perceived as one-sided blends (i.e. blends where one input space is dominant), since categorical perception seems to play a significant role (especially in relation to the upward leading note movement). The creative harmonisation assistant described in the previous paragraph was shown to be able to express the harmonic character of diverse idioms in a creative manner, while the blended harmonies often extrapolated the two input idioms, creating novel harmonic concepts. The nature of the perceptual impact of the blended harmonisation products generated by the system was examined in [17]. In this work, the behavioural assessment of system-generated blended harmonisations revealed that the system has succeeded in producing perceivable blends – both across idioms, modes and types of chromaticism – that were equally preferred, compared to non-blends.

This deliverable consists of the above cited papers, which are given in the following pages.

References

Full texts of all the publications produced in the context of this deliverable, given in the list of references that follows, are available on the COINVENT webpage. Some indicative publications that are shown in bold, which are representative of the work done on theoretical aspects of blending in music, harmonic learning, harmonic blending and evaluation of harmonic blending products, are attached in this deliverable.

[1] CAMBOUROPOULOS, E. The Harmonic Musical Surface and Two Novel Chord Representation Schemes. In *Computational Music Analysis*, D. Meredith, Ed. Springer, 2015, pp. 31–56.

[2] CAMBOUROPOULOS, E., KALIAKATSOS-PAPAKOSTAS, M., AND 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)* (Manchester, UK, 2015).

[3] EPPE, M., CONFALONIER, R., MACLEAN, E., KALIAKATSOS-PAPAKOSTAS, M., CAMBOUROPOULOS, E., SCHORLEMMER, M., CODESCU, M., AND KÜHNBERGER, K.-U. Computational invention of cadences and chord progressions by conceptual chord-blending. In *International Joint Conference on Artificial Intelligence (IJCAI) 2015* (2015).

[4] KALIAKATSOS-PAPAKOSTAS, M., CONFALONIER, R., CORNELI, J., ZACHARAKIS, A., AND CAMBOUROPOULOS, E. An argument-based creative assistant for harmonic blending. In *Proceedings of the 7th International Conference on Computational Creativity (ICCC)* (2016).

[5] **KALIAKATSOS-PAPAKOSTAS, M., MAKRIS, D., TSOUGRAS, C., AND CAMBOUROPOULOS, E. Learning and creating novel harmonies in diverse musical idioms: An adaptive modular melodic harmonisation system. *Journal of Creative Music Systems 1*, 1 (2016)**.

[6] KALIAKATSOS-PAPAKOSTAS, M., MAKRIS, D., ZACHARAKIS, A., TSOUGRAS, C., AND CAMBOUROPOULOS, E. Learning and blending harmonies in the context of a melodic harmonisation assistant. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (2016), Springer, pp. 520–527.

[7] **KALIAKATSOS-PAPAKOSTAS, M., QUEIROZ, M., TSOUGRAS, C., AND CAMBOUROPOULOS, E. Conceptual blending of harmonic spaces for creative melodic harmonisation. *Journal of New Music Research* (2016), (Submitted – Under review).**

[8] KALIAKATSOS-PAPAKOSTAS, M., ZACHARAKIS, A., TSOUGRAS, C., AND 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).

[9] KALIAKATSOS-PAPAKOSTAS, M., ZACHARAKIS, A., TSOUGRAS, C., AND CAMBOUROPOULOS, E. Modelling cadence perception via musical parameter tuning to perceptual data. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (2016), Springer, pp. 552–561.

[10] MAKRIS, D., KALIAKATSOS-PAPAKOSTAS, M., AND CAMBOUROPOULOS, E. A probabilistic approach to determining bass voice leading in melodic harmonisation. In *Mathematics and Computation in Music: Proceedings of the 5th International Conference, MCM 2015* (London, UK, June 2015), Springer, Berlin.

[11] MAKRIS, D., KALIAKATSOS-PAPAKOSTAS, M., AND 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).

[12] STEFANOU, D. Investigating Social Creativity and Concept Invention in Collaborative Musical Situations. In *Proceedings of the First International Conference on New Music Concepts* (Treviso, Italy, 2015).

[13] STEFANOU, D., AND CAMBOUROPOULOS, E. Enriching the blend: Creative extensions to conceptual blending in music. In *Proceedings of the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM)* (Manchester, UK, 2015).

[14] **TSOUGRAS, C., AND STEFANOU, D. Conceptual Blending and Meaning Construction: A Structural/hermeneutical Analysis of the 'Old Castle' from Musorgsky's 'Pictures at an Exhibition'**. In ***Proceedings of the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM)*** **(Manchester, UK, 2015)**.

[15] ZACHARAKIS, A., KALIAKATSOS-PAPAKOSTAS, M., AND CAMBOUROPOULOS, E. Conceptual blending in music cadences: A formal model and subjective evaluation. In *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2015)* (Malaga, Spain, 2015).

[16] **ZACHARAKIS, A., KALIAKATSOS-PAPAKOSTAS, M., TSOUGRAS, C., AND CAMBOUROPOULOS, E. Creating musical cadences via conceptual blending: Empirical evaluation and enhancement of a formal model. *Music Perception* (2016), (Submitted – Under revision).**

[17] **ZACHARAKIS, A., KALIAKATSOS-PAPAKOSTAS, M., TSOUGRAS, C., AND CAMPOUROPOULOS, E. Empirical methods for evaluating music structural blends: a case study on melodic harmonisation. *Musicae Scientiae* (2017), (forthcoming)**.

Proceedings of the Ninth Triennial Conference of the European Society for the Cognitive Sciences of Music, 17-22 August 2015, Manchester, UK
Ginsborg, J., Lamont, A., Phillips, M., Bramley, S. (Editors)

Conceptual blending and meaning construction: A structural/hermeneutical analysis of the 'Old Castle' from Musorgsky's 'Pictures at an Exhibition'

Costas Tsougras,[1] Danae Stefanou[2]
School of Music Studies, Aristotle University of Thessaloniki, Greece
[1]tsougras@mus.auth.gr, [2]dstefano@mus.auth.gr

ABSTRACT

Conceptual blending is a cognitive theory proposing the combination of diverse conceptual spaces for the creation of novel blended spaces. Musical conceptual blending can be intra-musical, pertaining to the combination of diverse structural elements for the creation of new melodies, harmonies or textures, as well as cross-domain, involving the integration of musical and non-musical spaces for the creation of novel analogies or metaphors. The present paper presents a structural and hermeneutical analysis of 'Il vecchio castello' from Modest Musorgsky's 'Pictures at an Exhibition' in an attempt to disclose both the intra-musical (combination of modal, tonal and coloristic harmonic spaces) and the extra-musical (contextual, symbolic and programmatic aspects) conceptual blending that the work incorporates. The analysis reveals that the piece comprises seven strophes of a song form that emerge from a common melodic core, through the dynamic evolution of harmonic spaces from diatonic modality to impressionistic/coloristic chromaticism and with the combinatorial use of ten harmonization concepts. The reductional/prolongational analysis provides input for two distinct Conceptual Integration Networks, the first describing the intra-musical blending of melodic harmonization and the second proposing the cross-domain blending of the musical and pictorial input spaces into a blended hermeneutical space that projects the work's narrative/programmatic/emotional potential. The proposed analysis shows how musical structure promotes meaning construction through cross-domain mapping. This research suggests that conceptual blending theory as an analytical tool can promote a richer structural interpretation and experience of Musorgsky's work.

I. Introduction

From a traditional musico-analytical perspective, Musorgsky's 'Pictures at an Exhibition' is a typical example of programme music. It refers to a series of paintings, and the imaginary affective exploration of their features. This programme, in keeping with 19th-century formalist distinctions between intrinsically musical features and extra-musical interpretations attached to them, is seen as somehow secondary and 'parallel' to the music.

In this paper we argue for a somewhat different interpretation, drawing on the theory of conceptual blending (Fauconnier & Turner 2003) and related work on metaphor & cross-domain mapping (e.g. Zbikowski 2002 & 2008, Spitzer 2003). Through a case-study analysis of the 'Old Castle', we explore instances of conceptual blending which go beyond the idea of a programme that is merely applied onto the musical work, and re-cast Musorgsky's composition as a dynamic, multiple-level integration of incongruous temporal, spatial and affective modalities.

A fundamental assumption for this investigation is the idea of a scored composition as an emergent structure, which can also be studied retrospectively. The ensuing analysis is therefore intended to provide a possible interpretation of how we listen to the 'Old Castle', how this process generates meaning that is neither purely musical nor exclusively pictorial or verbal, and how the elements that are central to this blended understanding of the work, are arguably themselves a result of structural blending.

A. Perspectives from Cognition and Philosophy of Mind: Conceptual Blending and Qualia

Fauconnier & Turner's Conceptual Blending theory (2003) is a step further from unidirectional theories of metaphor, most notably Lakoff & Johnson's (1980) Conceptual Metaphor Theory (CMT). CMT suggests that we map concepts across different domains, borrowing features from one source (e.g. painting) and applying them to a target (e.g. music), so that the attributes of the source domain are mapped onto those of the target (e.g. 'nuanced dynamics' or 'a dark tonality'). Blending, on the other hand, presupposes an equilateral, multi-directional relationship not only between different domains, but between *conceptual spaces*. These spaces may be contrastive or qualitatively different, and may only share some structural features between them. In that sense, we may also identify blends situated exclusively within the domain of music, e.g. between clashing chords or contrasting tonalities (Ox 2014, Kaliakatsos-Papakostas *et al* 2014), as well as blends combining properties of text, image and sound, e.g. in cinema or advertising (Cook 2001) or in recorded pop songs (Moore 2012).

Applications of conceptual blending in music analysis are still relatively few and rather general in nature. Cook (2001, see also an earlier attempt in Cook 1998) makes one of the first attempts to represent a music and moving image blend in his analysis of a Citroen car commercial, while Zbikowski (2002) provides one of the more detailed analyses to date of how text painting and programme music operate together on the basis of conceptual blending. While more recent authors (e.g. Schmidt 2012) have also proposed critical re-examinations of these analytical approaches, for the purposes of this paper, we will primarily rely on Zbikowski's paradigm, not only because it is the most analytically inclined example of current literature on blending in music, but also because of its closeness to the material under study (a complex programmatic work involving several layers of visualisation and meaning construction).

As 'the work' in this case is not merely a musical text, and the composer's relationship to the source material is more complex than the kind of one-directional representation or *ekphrasis* suggested e.g. by Bruhn (2000), we also refer to the qualitatively different, contrastive states that account for the piece's multiple dimensions as *qualia*. Though the properties of qualia have been the subject of extended criticism among consciousness theorists (most notably Dennett 1991), the idea of otherwise indescribable differences in consciousness between past and present, reality and dream, depiction and interpretation etc. is a useful way to conceptualize the deeper-level structures that permeate the composition.

B. Musorgsky's 'Pictures' and the 'Old Castle'

Musorgsky's 'Pictures at an Exhibition' is a well-known piano suite, inspired from paintings and architectural drawings by Viktor Alexandrovich Hartmann (1834-1873), a close friend of the composer, put on display during a posthumous exhibition of 400 of his works in February and March 1874 in St. Petersburg. The suite comprises 10 pieces and 5 promenades that function as preludes and/or bridges. It was written in one creative burst in June 1874 (Russ 1992; Oldani; Brown 2002: 229-241).

The suite, according to Russ (1992: ch. 1; see also Taruskin 2010: vol. 3, ch. 12) incorporates Musorgsky's key stylistic elements: nationalism, populism, anti-romantic realism and conscious distance from mainstream (Germanic) concepts of musical form, motivic development and harmonic structure. A narrative dimension has been identified and commented upon in all of Musorgsky's 'musical pictures' (Russ 1992: 31, Tarasti 1994: ch. 8), as if the composer focuses on someone or something within the picture and creates a story about it through music, thus forging an indivisible duality of psychological state/musical structure for each piece.

The 'Old Castle' is the second piece of the suite, entitled by the composer in Italian as 'Il vecchio castello'. The original watercolor painting has been lost or sold during the exhibition (Brown 2002: 230), but according to Stasov's description (Frankenstein 1939: 282), it was a depiction of "a medieval castle, before which stands a singing troubadour". Bibliographic references to the piece (Russ 1992: 37-38; Tarasti 1994: 214, 227-229; Brown 2002: 235) stress its modal Russian character, its siciliana rhythmic pattern and its melancholic mood, but do not include full or partial musical analysis. We cite two of these references, since they indirectly reflect the present analytical approach (italics by the authors): Eero Tarasti, in his semiotic analysis (1994: 214), refers to the piece as Italian pastiche, where "the 'old castle' alludes to the past, a *heterotopic* place, 'elsewhere' with respect to the musical narration", and David Brown mentions (2002: 235) that "the melody that runs throughout the piece is his [the minstrel's] song, a *blend* of Italian siciliana with Russian melancholy".

C. Research Aims

In Conceptual Blending Theory (Fauconnier and Turner, 2003), elements from diverse, but structurally related, mental spaces are 'blended', giving rise to new conceptual spaces that often possess new powerful interpretative properties, allowing better understanding of known concepts or the emergence of novel concepts. Conceptual blending allows the construction of meaning by correlating elements and structures from diverse conceptual spaces.

The present research's aim is the exploration of conceptual blending between the musical and pictorial spaces embedded in the 'Old Castle'. The inquiry was triggered by the piece's implicit heterogeneity regarding its modal/tonal content, a feature that seems to grow and expand while the music evolves from beginning to end, while constantly revolving around a stable rhythmic pattern and a common melodic core. The analysis will therefore pursue an explication of the multi-directional metaphoric relation between music and picture through structural music analysis and cross-domain mapping, as well as a description of its dynamic evolution. For this purpose, multi-level ontologies in music will be employed in multi-level blending through the basic operations of composition, completion and elaboration (Zbikowski 2002: 80).

II. Music Analysis

The piece can be considered, in accordance with Russ's description (1992: 37), an Italian "serenade ... [that] turns into a Russian song without words", where a diatonic modal melodic core unfolds differently in each stanza, evoking different harmonizations. The analysis that follows focuses on harmonic and prolongational structure, making references to rhythmic and textural aspects. Our choice for using prolongational analysis and revealing quasi-Ursatz schemas may seem at first inappropriate for music that consciously avoided mainstream harmonic and developmental theories and practices (Russ 1992: 9). However, the specific piece affords the application of such a methodology, albeit in an idiomatic way, due to its linear texture (see also Puffett 1990 and Russ 1990 for prolongational analyses of 'Catacombs' and 'Nursery').

A. Form and compositional concepts

The piece is in strophic song form, with a short introduction and seven stanzas of unequal length, as shown below:

Introduction (b. 1-8)
Stanza 1 (b. 9-18)
Stanza 2 (b. 19-28)
Stanza 3 (b. 29-37)
Stanza 4 (b. 38-50)
Stanza 5 (b. 51-69)
Stanza 6 (b. 70-95)
Stanza 7 (b. 96-107)

Ten main compositional concepts can be identified, employed in various combinations by Musorgsky for the composition of the seven stanzas:

1. Drone of tonic in the lower voice (omnipresent and constant throughout the whole piece)
2. Siciliana rhythmic pattern and subpatterns: ♩ ♪ ♩ ♪ | ♪. ♬ ♫♩, etc (the main pattern for the tonic drone and related subpatterns for the melody)
3. Diatonic modal harmony (diatonic voice-leading, free non-functional use of triads for melodic harmonisation in the context of the diatonic modes)

4. Diatonic tonal harmony (functional use of chords for melodic harmonisation in the context of major-minor tonality, diatonic voice-leading, tonal cadence schema: iv-V^7-i)

5. Chromatic tonal harmony (use of more dissonant chords, chromatic mixture, tonicizations, chromatic voice leading)

6. Chromatic coloristic/impressionistic harmony (free use of chromatic sonorities without tonal harmonic function)

7. Modal interchange (change of mode while keeping the same pitch center) and hyper-modulation (change of pitch tonal space)

8. Parallel harmony (diatonic or chromatic/real chord planing)

9. Scale of sensory dissonance (conscious use of intrinsic dissonance level for the choice of chords)

10. Fragmentation of musical texture (use of unconnected snippets / mosaic texture)

These concepts can be categorized –with categorical overlapping– as rhythmic (1, 2, 10), harmonic (1, 3, 4, 5, 6, 7, 8), textural (8, 10) and cognitive/schematic (9, 10).

B. Analysis of the seven stanzas

In this subsection an analysis of each stanza is presented, focusing on the compositional concepts employed and illustrated with two-level prolongational graphs.

Introduction and Stanza 1 (b. 1-18). The left-hand introduction and the first melodic stanza are purely diatonic, with their pitch content coming from the G# Aeolian mode, and with characteristic descending voice-leading (5-4-3-2-1 for the intro and 8-7-6-5-4-3 for the melody). The intro segment is also repeated as a codetta (fig. 1).

The concepts employed are: tonic drone, siciliana rhythm, modal harmony (G# Aeolian, descending diatonic voice leading).

2. „IL VECCHIO CASTELLO."

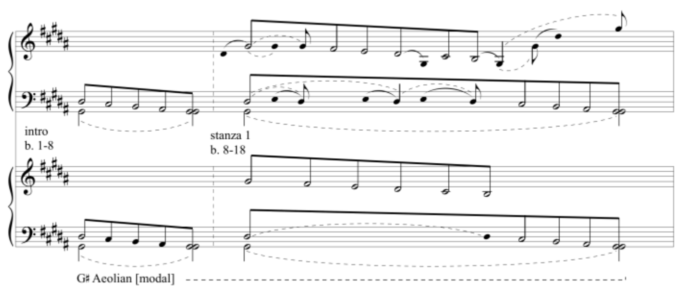


Figure 1. Score & prolongational analysis of Intro and Stanza 1.

Stanza 2 (b. 19-29). The second stanza starts similarly in the G# Aeolian mode, but 3 bars later the use of A natural denotes a modal interchange towards the G# Phrygian. The parallel $^6_3$ chords that introduced the modal interchange continue, creating a tonicization of the C# minor chord. This is subsequently used as a iv harmonic degree in G# minor tonality, leading to a V^7-i (fig. 2). Thus, although the main melodic line is the same (8-7-6-5-4-3), a hyper-modulation from the modal to the tonal system occurs (fig. 2).

Concepts employed: tonic drone, siciliana rhythm, modal harmony (G# Aeolian – G# Phrygian, descending voice-leading 8-7-6-5-4-3), tonal harmony (G# minor, cadence iv-V^7-i), modal interchange, hyper-modulation, parallel harmony (diatonic $^6_3$ chords).

Figure 2. Score & prolongational analysis of Stanza 2.

Stanzas 3, 4. (b. 29-37 & 38-50). The exploration of diatonic modes based on G# continues in these two almost identical stanzas (their only difference is that the fourth stanza includes the intro segment as a codetta). The stanza begins in G# Phrygian (A natural), interchanges to G# Locrian (A, D natural), returns to G# Aeolian and concludes in G# minor tonality. The expanded modal interchange concept introduces a mode not used in the Middle Ages, the Locrian, conveying a more Russian/19th-century profile to the stanza's modality (fig. 3).

Concepts employed: tonic drone, siciliana rhythm, modal harmony (G# Phrygian – G# Locrian – G# Aeolian, descending voice-leading 6-5-4-3), tonal harmony (G# minor, cadence iv-V^7-i), modal interchange, hyper-modulation.

Figure 3. Score & prolongational analysis of Stanzas 3 and 4.

Stanzas 5, 6 (b. 51-69 & 70-95). Stanza 5 begins in G# Aeolian, but then, when the melody ascends chromatically from G# to D#, chromatic harmony is employed for its harmonization. Initially, two tonicizations take place in A# major and C# major (through secondary diminished 7th chords). Subsequently, the two last melodic steps (Cx-D#) are harmonized with intrinsically dissonant non-functional chromatic sonorities (D#-F#-A#-Cx, E#-G#-D#), before reaching C# minor through an embellishing non-functional chord (E-G#-Cx), and finally arriving at a functional stable harmonization of D# (D# major chord). These non-functional coloristic/impressionistic chords have diminishing sensory dissonance levels, a parameter exploited by the composer in the transition from tension to relaxation: [D#-F#-A#-Cx] - [E#-G#-D#] - [E-G#-D] - [E-G#-C#]. The stanza closes with a cadence to G# minor tonality (iv-V^7-i), that also completes the background melodic voice-leading (5-4-3). This stanza greatly expands the concept of hyper-modulation, incorporating four distinct harmonic systems (modal, diatonic tonal, chromatic tonal, impressionistic), each pertaining to a different tonal pitch space / historical era (fig. 4).

Figure 4. Score & prolongational analysis of Stanzas 5 and 6.

Stanza 6 is almost identical, but with an extra element: the fragmentation of the musical texture by employing snippets of the previous stanzas (b. 87-95), having as a result the absence of the cadential pattern V^7-i at its end: the unresolved V^7 of b. 86 is prolonged until b. 95 (fig. 5).

Figure 5. Reductional analysis of Stanza 6.

Stanzas 5 and 6 incorporate almost all the compositional concepts: tonic drone, siciliana rhythm, modal harmony (G# Aeolian), chromatic tonal harmony (vii^{o7}-I, chromatic ascending voice leading, brief tonicizations), coloristic harmony (D#m^{7M}-E#$m^{7/-5}$), diatonic tonal harmony (G# minor, cadence iv-V^7-i), hyper-modulation, sensory dissonance scale, fragmentation.

Stanza 7 (b. 96-107). The last stanza returns to the initial melodic material, albeit with more chromaticism (chromatic voice-leading, altered diminished 7th chord for the tonicization of iv). Michael Russ (1992: 38) argues that this is a coda, but we will disagree, because this part contains the structural ending of the work, the only complete iteration of the piece's melodic core: the descending voice-leading schema (8-7-6-5-4-3-2-1) (fig. 6).

Concepts employed: tonic drone, siciliana rhythm, modal harmony (G# Aeolian), chromatic tonal harmony (descending chromatic voice-leading, altered chords), diatonic tonal harmony (G# minor, iv-V^7-i), perfect cadence with structural closure, hyper-modulation.

Figure 6. Score & prolongational analysis of Stanza 7.

C. Summary of compositional features

The preceding musical analysis has revealed that the 'Old Castle' is essentially the result of seven different evolutions of a common modal melodic core –namely a descending voice-leading linear structure–, through the dynamic evolution of harmonic spaces from diatonic modality to diatonic/chromatic tonality and impressionistic/coloristic chromaticism, with the combinatorial use of ten compositional concepts. The harmonic evolution is supported by the omni-present common element of the siciliana tonic drone, and occurs linearly, starting with diatonic modality in the 1st stanza, culminating with the use of all four spaces in the 6th stanza (through hyper-modulations) and closing with the tonal cadence in the 7th stanza and the completion of the melodic schema.

III. Conceptual Integration Networks

This section attempts to put the analysis results in context, drawing on Zbikowski's representation of conceptual blending in music. So, two different Conceptual Integration Networks (CINs) will be constructed, each with its own generic, input and blended spaces, and with reference to Fauconnier & Turner's (2003) typology of single-scope and double-scope blending networks.

A. "Intra-musical" structural blending

CIN 1 (Conceptual Integration Network 1) proposes that the piece's evolutionary musical structure is a result of the intra-musical blending of harmonic spaces through the concept of hyper-modulation. So, the Generic Space, to which both input spaces relate, is *Music-Song*; it is defined by parameters of melody, rhythm, harmony, hierarchy and musical texture. Input Space 1 is *Melody* (properties: modes/scales, structural pitches, melodic/linear cadences, interval succession/voice-leading, implied harmony, rhythm) and Input Space 2 is *Harmony* (properties: diatonic modality, diatonic tonality, chromatic tonality, coloristic harmony, hyper-modulation, parallel harmony, pedal notes/drones, harmonic rhythm). The combinations that the two input spaces afford yield the Blended Space, i.e. the musical structure of '*Il vecchio castello*', as an evolutionary succession of seven different melody/harmony amalgams produced by the combination of four harmonic spaces (fig. 7).

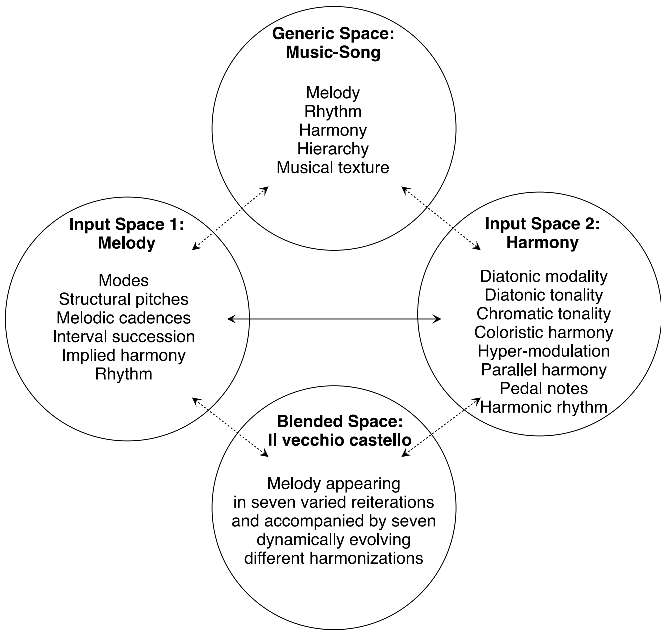


Figure 7. CIN 1: "Intra-musical" structural blending.

B. Cross-domain conceptual blending (meaning construction)

CIN 1 could be seen along the lines of Fauconnier and Turner's (2003) single-scope blending, where the re-framing of a concept (melody) through a different set of relations (harmony) results in changing instantiations of the concept. CIN 2 (Conceptual Integration Network 2) proposes a double-scope blending of the musical and pictorial input spaces into an integrated conceptual space, which projects the work's narrative and emotional potential and further promotes novel meaning construction. As Turner (2003) notes, double-scope blending is one of the most creative cognitive features associated not only with the conceptualization of everyday reality, but particularly with the formulation of artistic and scientific concepts. Double-scope networks involve the simultaneous elaboration of two contrasting input spaces, and the running of two previously unrelated scripts as one blend. Being in one place, in one time, and fully perceiving and interacting with the features of this place and time, while also simultaneously recollecting and exploring another place, at another time, is a typical example of double-scope blending.

The Generic Space for CIN 2 involves *Contrasting Ontological States*, and it can be split into four contrasting generic sub-spaces: *Temporality*, *Spatiality*, *Affect* and *Qualia*, each producing a separate sub-CIN. Input Space 1 is the *Pictorial Space*, Input Space 2 is the *Musical Space* (or one of its constituents), and the Blended Space is '*Il vecchio castello*' as a perceived programmatic musical work.

CIN 2a: *Contrasting Temporality* (fig. 8). This CIN describes the contrasting temporality embedded in the piece, as a result of the contrasting harmonic spaces employed and the contrasting epochs they correspond to in the pictorial space (contrast between the depiction of the medieval castle in the past and its reception in a 19th-century 'present').

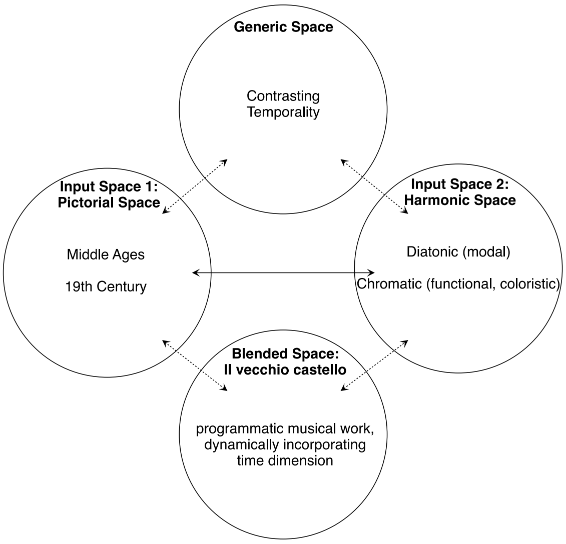


Figure 8. CIN 2a: Cross-domain blending - Temporality.

CIN 2b: Contrasting Spatiality (Geographic/national marker). This CIN (fig. 9) describes the embedded contrasting spatiality, expressed at the pictorial space by the depiction of an Italian castle observed in a Russian gallery and at the musical/melodic space by an Italian siciliana melody/rhythm implanted with Russian folk character and corresponding modality. Moreover, the Italian element is declared in Musorgsky's original Italian title, and the Russian vernacular element has been associated with a type of melismatic peasant song known as *protyazhnaya* (Russ 1992: 51).

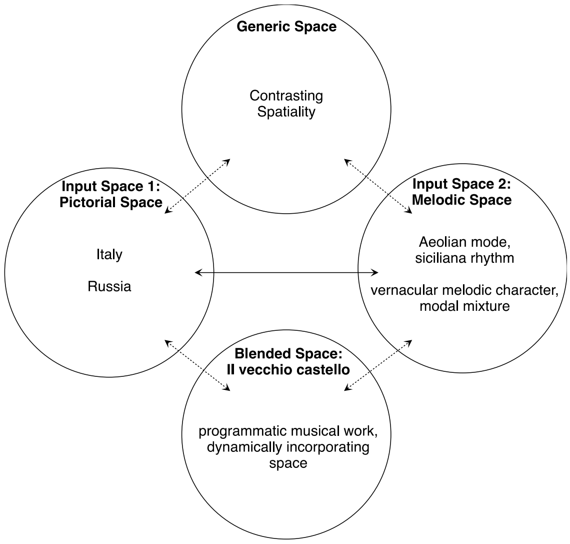


Figure 9. CIN 2b: Cross-domain blending - Spatiality.

CIN 2c: *Contrasting Affective States* (emotion). This CIN (fig. 10) describes the contrasting affects (emotions) that may be evoked by the blending of the pictorial and musical input spaces. 'Love' (expressed in pictorial space by the singing troubadour) can be experienced as 'Nostalgia for love', under the influence of the musical space, where a serenade gradually turns into a melancholic folk song.

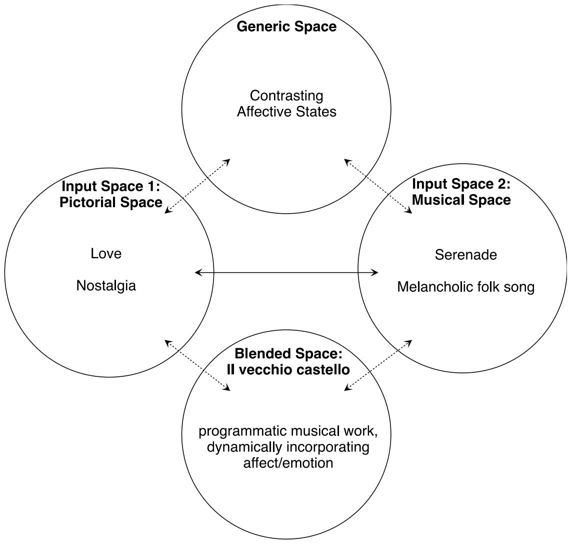


Figure 10. CIN 2c: Cross-domain blending - Affect.

CIN 2d: *Contrasting Qualia*. This CIN describes the different instances of subjective, conscious experience (formulated as *qualia,* after Goguen 2004) embedded in the music in latent form. The contrasting qualia, in this case, refer to two different kinds of psychological/consciousness states, which can be inferred in input Space 1 (pictorial). They are the state of real-time conciousness, and the state of dream/fading recollection, corresponding to the idealized "real" past and the imaginary "dreamy" present. These states are reflected in Input Space 2 (musical) as the juxtaposition of normal rhythmic flow of melody/form and fragmentary array of snippets or the contrast of simple strophic and dynamically evolving song form.

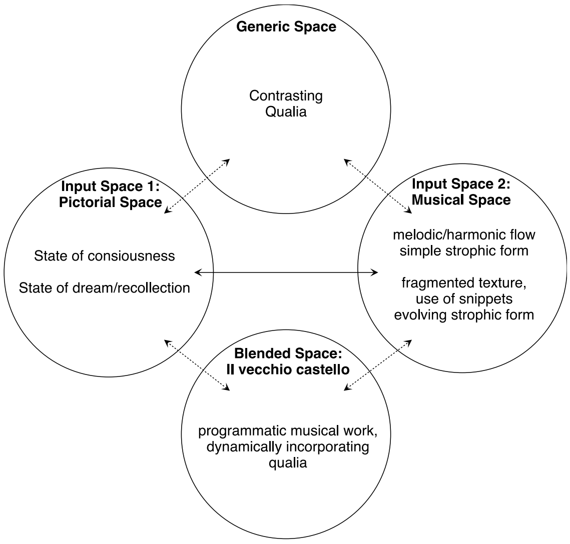
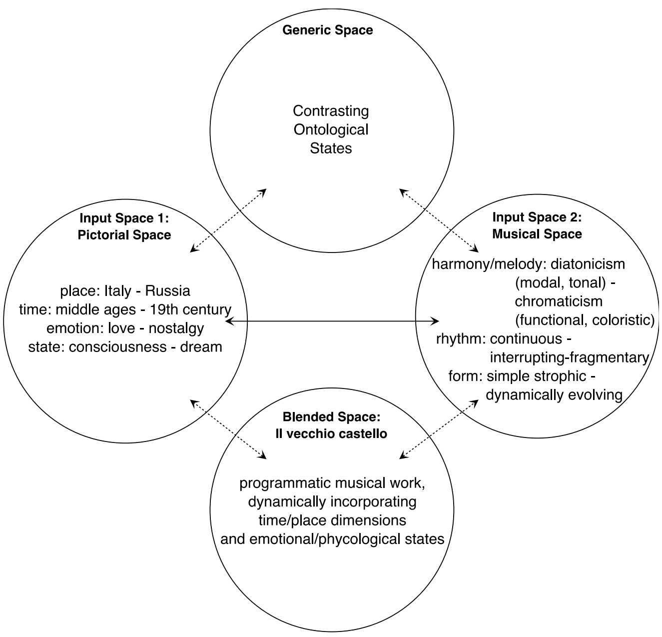


Figure 11. CIN 2d: Cross-domain blending - Qualia.

Overall, CIN 2 (Conceptual Integration Network 2) proposes meaning construction through double-scope conceptual blending and emerges as the union of its four constituent sub-CINs described above. This collective, multiple-scope, multiple-level CIN suggests that the contrasting ontologies embedded in the musical structure trigger contrasting ontologies in the projected "perceived/imagined" pictorial space, and that this cross-domain integration elicits a richer aesthetic experience for the listener.



C. Figure 12. CIN 2: Cross-domain blending.

D. Dynamic evolution

Moreover, a dynamic evolution of conceptual blending takes place as the piece progresses from the first stanza to the last, as if following a narrative path, through which the "real", representational drawing of the Italian castle with the love-singing troubador gradually becomes a "dreamy" abstraction of an old castle, vaguely remembered and evoked in another time and place. This occurs due to the blending operation of *elaboration*, which denotes an imagination-triggering process that stems from musical structure and constructs emergent emotions and meaning (fig. 13).

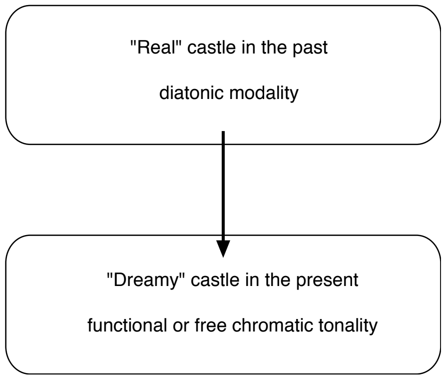


Figure 13. Dynamic evolution of conceptual blending.

IV. Meaning Construction - Conclusions

Conceptual blending in this case involves the use of harmonic, melodic, formal, textural and schematic elements that are not compatible with a simple depiction of a medieval castle. Through blending and cross-domain mapping, music precipitates the listener to "see" or imagine the castle gradually lost into the vortex of time, misty, dreamy, in an obscure place, and with the feeling of chivalrous love gradually transformed into melancholic nostalgia as the music unfolds.

Consequently, the *old castle* that one might see in the painting is very different from the '*old castle*' that our imagination creates while experiencing Musorgsky's piece, and this transcendence to a much richer aesthetic experience is feasible through the blending of the pictorial and musical conceptual spaces.

In effect, as we move from simple cross-domain mapping between music and image, onto the single-scope binding of melody and harmony (CIN1) and higher-level, double-scope blending functions (CIN2), it is possible even to explore the work as a process of cognitive integration (between melodic and harmonic elements, visual and auditory references) and dis-integration between contrastive, qualitatively different temporal, spatial and affective states. According to Bache (2005) dis-integration is one of the most important features of higher-level blending. We elaborate and "make sense of" blends only by consciously focusing on the differences between input spaces and thus acknowledging the terms on which a metaphor operates.

A present-day listener is thus able to conceive of Mussorgsky's 'Old Castle' as an imaginary castle, a wordless song, a nostalgic reverie, a musical landscape, or all of these at once. This begs a bigger question regarding the levels of mediation (Stefanou 2004) involved in this metaphorical concept construction, from Hartmann's sketches up to Mussorgsky's score, and even more so, a performance of it. Further extensions of the present research could engage with the dimension of performance, and its role in the complex blending procedures suggested here. While it has not been possible to do so within the limited confines of this research, a focus on performance and listening would probably significantly enrich the Conceptual Integration Networks proposed above, and also help situate the analysis in terms of embodied meaning.

Finally, a broader issue could be raised by the very conceptualization of the work's features and the choice to represent them in two distinct types of networks. By distinguishing intra-musical from cross-domain conceptual blends, we do not wish to imply that meaning and structure are exclusively associated with one space or other. Quite on the contrary, we think that CIN1 and CIN2 could themselves become part of a multiple-scope blend, exposed by this categorization, and involving so-called intra-musical and extra-musical features. This separation is in itself the result of a conceptual metaphor (Spitzer 2003), by which "music" is equated with structure, and seen as a central locus, outside of which various other domains are tangentially involved in the production of secondary meaning. Hopefully, in this research we have also opened up a space for further problematization and relativization of the conceptual metaphor of intra- and extra-musicality, and further research can elucidate the particular terms on which it operates.

ACKNOWLEDGMENT

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.

REFERENCES

Bache, C. (2005). Constraining Conceptual Integration Theory: Levels of Blending and Disintegration. *Journal of Pragmatics*, 37, 1615-1635.

Brown, D. (2002). *Musorgsky: His Life and Works*. Oxford: Oxford University Press.

Bruhn, S. (2000). *Musical Ekphrasis: Composers Responding to Poetry and Painting*. New York: Pendragon.

Cook, N. (1998). *Analysing Musical Multimedia*. Oxford: Oxford University Press.

Cook, N. (2001). Theorizing musical meaning. *Music Theory Spectrum*, 23/2, 170-195.

Dennett, D. C. (1991). *Conciousness Explained*. Boston: Little, Brown and Co.

Fauconnier, G. & Turner, M. (2003). *The Way We Think: Conceptual Blending And The Mind's Hidden Complexities*. New York: Basic Books [reprint edition].

Frankenstein, A. (1939). Victor Hartmann and Modeste Musorgsky. *The Musical Quarterly*, 25/3, 268-291.

Goguen, J. A. (2004). Musical Qualia, Context, Time and Emotion. *Journal of Consciousness Studies*, 11/3-4, 117-147.

Kaliakatsos-Papakostas, M.; Cambouropoulos, E.; Kuhnberger, K.U.; Kutz, O.; Smaill, A. (2014). Concept Invention and Music: Creating Novel Harmonies via Conceptual Blending. *Proceedings of the 9th Conference in Interdisciplinary Musicology – CIM14*.

Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Moore, A. F. (2012). *Song means: Analysing and interpreting recorded popular song*. Aldershot: Ashgate.

Oldani, R. W. "Musorgsky, Modest Petrovich". Entry in *Grove Music Online* <accessed 10/5/2015>.

Ox, J. (2014). Analogy and Conceptual Blending are Part of a Visualization Toolkit for Artists and Scientists: Introducing the Cognitive Space Transfer. *Proceedings of the IEEE VIS2014 Arts Program VISAP'14: Art + Interpretation.* Paris, France, November 2014.

Puffet, D. (1990). A Graphic Analysis of Musorgsky's 'Catacombs'. *Music Analysis*, 9/1, 67-77.

Russ, M. (1990). The Mysterious Thread in Musorgsky's 'Nursery'. *Music Analysis*, 9/1, 47-65.

Russ, M. (1992). *Musorgsky: Pictures at an Exhibition* (Cambridge music handbooks). Cambridge: Cambridge University Press.

Spitzer, M. (2003). *Metaphor and Musical Thought*. Chicago: University of Chicago Press.

Schmidt, S. (2012). Some Remarks Concerning the Generative Theory of Tonal Music, Conceptual Metaphors and Conceptual Blending in Music <retrieved from www.researchgate.net>.

Stefanou, D. (2004). Placing the Musical Landscape: Performance, Spatiality, and the Primacy of Experience. PhD thesis, University of London. <retrieved from http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.406939>

Tarasti, E. (1994). *A Theory of Musical Semiotics*. Bloomington: Indiana University Press.

Taruskin, R. (2010). *Oxford History of Western Music*. New York: Oxford University Press.

Turner, M. (2003). Double-Scope Stories. In D. Herman (ed.) *Narrative Theory and the Cognitive Sciences*. Stanford: CSLI Publications.

Zbikowski, L. M. (2002). *Conceptualizing Music*. New York: Oxford University Press.

Zbikowski, L. M. (2008). Metaphor and Music. In R.W. Gibbs (ed.) *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press.

Learning and Creating Novel Harmonies in Diverse Musical Idioms: An Adaptive Modular Melodic Harmonisation System

Maximos Kaliakatsos-Papakostas
Department of Music Studies,
Aristotle University of Thessaloniki,
54124, Thessaloniki, Greece.
email: maxk@mus.auth.gr

Dimos Makris
Department of Informatics Studies,
Ionian University,
49100, Corfu, Greece.
email: c12makr@ionio.gr

Costas Tsougras
Department of Music Studies,
Aristotle University of Thessaloniki,
54124, Thessaloniki, Greece.
email: tsougras@mus.auth.gr

Emilios Cambouropoulos
Department of Music Studies,
Aristotle University of Thessaloniki,
54124, Thessaloniki, Greece.
email: emilios@mus.auth.gr

Abstract

Melodic harmonisation is a sophisticated creative process that involves deep music understanding and specialised music knowledge relating to melodic structure, harmony, rhythm, texture, form. In this paper a new melodic harmonisation assistant is presented that is adaptive (learns from data), general (can cope with any tonal or non-tonal harmonic idiom) and modular (learns different aspects of harmonic structure such as chord types, chord transitions, cadences and voice-leading). This melodic harmonisation system can be used, not only to mimic given harmonic styles, but, to generate novel harmonisations for diverse melodies by exploring the harmonic possibilities provided by the implied harmonies of input melodies, or by allowing the imposition of user-defined chord constraints leading thus to new unforeseen harmonic realisations. The various components of the proposed model are explained, and, then, a number of creative harmonisations of different melodies are presented, along with an intuitive statistical analysis, to illustrate the potential of the system.

1 Introduction

Creative music systems are often criticised as not 'really' being creative *per se*; underlying this criticism is the belief that the actual human programmer is the true source of creativity. However, machine learning has made such criticisms more difficult to maintain, as a machine may acquire knowledge from data, construct a new conceptual space (new structural relations or even new elements) or explore an existing one without human intervention and, then, create new unforeseen output (Wiggins et al. 2009). Adaptability, flexibility, independence, autonomy are features associated with creativity (see key components of creativity in Jordanous (2013)); general representations and machine learning techniques allow creative systems to be open to new environments, to evolve, to transform existing or construct new concepts, to create new unexpected results.

A model of creativity has been proposed by Boden (2009) whereby a conceptual space may be explored by an agent in order to generate new creative instances (exploratory creativity) or the rules of the conceptual space are transformed changing the conceptual space itself (transformational creativity) or different conceptual spaces that share structural similarities are combined to create new blended spaces (combinational creativity). In the current study, the conceptual spaces are learned in a bottom-up fashion from data and are structured in a modular way, so as to allow (at a later stage) to combine different modules from different spaces creating thus new blended spaces. At this stage, the system is indicated to exhibit exploratory creativity, by composing harmonies that potentially excess the harmonic borders of a corpus.

This paper focuses on melodic harmonisation seen as a creative musical act. Some researchers follow the knowledge-engineering approach whereby experts encode the essential rules for harmonisation in a certain tonal style (from the Bach chorale expert system by Ebcioglu (1988) to the explicitly structured knowledge paradigm by Phon-Amnuaisuk and Wiggins (1999); Phon-Amnuaisuk et al. (2006)). In recent years, more attention has been given to probabilistic approaches that learn harmony from music data, using techniques such as Hidden Markov Models, N-grams, probabilistic grammars, inductive logic programming (Steedman (1996); Rohrmeier (2011); Conklin (2002); Scholz et al. (2009); Pérez-Sancho et al. (2009); Raphael and Stoddard (2004); Whorley et al. (2013); Dixon et al. (2010); Granroth-Wilding and Steedman (2014)). Such models automatically derive harmonic structure from training data and are thus more flexible than rule-based systems; however, they are applied usually to very narrow well-circumscribed tonal styles (e.g. Bach chorales or hymns or blues harmonies) and they generate acceptable harmonic progressions only within the corresponding learned harmony (cannot create new harmonies beyond the learned space).

This paper describes a creative melodic harmonisation system that can assist a user in generating new conventional or unconventional sequences of chords for a given melody. The creative melodic harmonisation assistant is based on a novel *General Chord Type* representation (Cambouropoulos et al. 2014) that allows the system to extract appropriate chord types from diverse harmonic idioms that comply with the traditional Western well tempered scale. For any harmonic idiom, the system learns from a set of pieces (harmonic reductions) the chord types that are relevant for the specific style, extracts probabilistic information on the most likely chord transitions (first-order transition tables), examines phrase endings with a view to establishing common endings/cadences, and learns basic features of voice leading (bass movement in relation to melodic motion, chord inversions and omission/duplication of chord notes). This information constitutes a conceptual space that characterises a specific harmonic idiom and is used to create original harmonisations for new previously unseen melodies.

Apart from learning harmony from a particular harmonic style and producing new harmonisations in this style, the current paper explores other creative aspects of the proposed melodic harmonisation assistant that diverge from a learned harmonic style (not automated in this phase). Firstly, a user may assign particular chords to specific melodic notes of a given melody, thus 'forcing' the system to explore harmonic regions of the learned harmonic space that are less common (or even alien) thus giving rise to potentially unexpected harmonisations, expressed as chord sequence paths that accommodate the selected chords. Secondly, a user may choose to harmonise a melody with different potentially incompatible learned harmonic styles (e.g. traditional folk melody with tonal harmony or octatonic harmony etc); potential inconsistencies are dealt with manually at this stage (automation of such processes is under development). The ultimate goal of this research is to enable a system to create original harmonisations by combining harmonic components of different harmonic spaces; such creative blending aspects are explored in Zacharakis et al. (2015); Cambouropoulos et al. (2015) and is part of ongoing research.

The proposed melodic harmonisation assistant is original in the following ways:

1. harmony is learned in an idiom-independent manner (i.e., harmonic features are acquired via machine learning for various tonal and non-tonal systems);

2. the system allows the exploration of a learned harmonic space by user-defined intermediate chords that may lead the system outside its expected course;

3. the creative system can use existing harmonic styles to harmonise melodies with 'incompatible' harmonic outlook.

In the following sections the proposed modular probabilistic melodic harmonisation system is presented; this system is able to learn different harmonic aspects (chord types, chord progressions, cadences, voice-leading) from practically any musical idiom and can use the acquired harmonic knowledge to harmonise novel melodies in innovative ways. The next section provides a short discussion of previous approaches to melodic harmonisation and an overview of the proposed system. Then, Section 4 analyses the module for constructing chord sequences by automatically employing cadences and allowing user-defined chord constraints. The module for fixing the voicing layout of chords is presented in Section 5 and finally several examples of melodic harmonisations in diverse harmonic idioms are given, along with an intuitive statistically-based analysis in Section 6.

2 Melodic harmonisation: related work and overview of the proposed system

Among the first approaches for capturing the characteristics of harmony in automated melodic harmonisation were ones that incorporated human expert knowledge (e.g. Ebcioglu (1988)) encoded in the form of rules, leading to systems that could harmonise melodies according to explicit stylistic directives. For a review of rule–based systems the reader is referred to Pachet and Roy (2001). A similar approach to rule–based methodologies is the one followed by systems that utilize genetic algorithms (GA), like the ones briefly reviewed by Donnelly and Sheppard (2011) and, also, in Phon-Amnuaisuk and Wiggins (1999). 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. The advantage of rule–based systems is that they can capture the hierarchical structure of complex musical idioms, e.g., by using grammar-related structures for tonal (Rohrmeier 2011; Koops et al. 2013) or especially focussed on jazz (Granroth-Wilding and Steedman 2014) music.

However, the melodic harmonisation methodologies that utilise rule-based techniques have a major drawback when dealing with melodic harmonisation in many diverse idioms: the encoding of rules that describe even a simple musical idiom is not always a realizable task, since idioms abound in complex and often contradicting interrelations between harmonic elements. In order to overcome such shortcomings, the formulation of *probabilistic* techniques and *statistical learning* has been proposed. Probabilistic techniques can, on the one hand, be trained on any musical idiom, given a set of harmonically annotated pieces, while on the other hand they encompass the possibility to take 'unusual' decisions if necessary – e.g. if the melody's implied harmony diverges from the learned harmony. Among many proposed methodologies, Bayesian networks (Suzuki 2013) and prediction by partial matching (Whorley et al. 2013) have been utilised 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 (Raczyński et al. 2013); and probabilistic graphical models for similar chord-assignment tasks (Paiement et al. 2006).

The main drawback of probabilistic methodologies, especially HMMs, is that they do not capture larger scale dependencies between remote harmonic parts (Pachet et al. 2011). For instance, phrase endings, indicated by cadences, are very distinctive parts of higher-level harmonic structure that are not captured by methodologies that concern chord-to-chord harmonic progressions. Cadences have been studied under different contexts in the computational harmonisation literature. For instance, in Borrel-Jensen and Hjortgaard Danielsen (2010), a methodology based on cadences was utilised to evaluate the outcomes of an automatic melodic harmonisation system. The methodologies presented in Allan and Williams (2004) and Hanlon and Ledlie (2002) utilise a backwards propagation of the hidden Markov model (HMM) methodology, starting from the end and constructing the chord progression in a backwards fashion, highlighting the role of the cadence in reflecting structure. In Yi and Goldsmith (2007) a probabilistic system was presented that rewarded those chord sequences that ended with a perfect cadence, while in Yogev and Lerch (2008) a probabilistic methodology that identifies the probable positions of cadences was introduced. Special consideration of cadences has also been employed in HMM-based methodologies, either by assigning an additional layer of probabilities for the final chords of sequences (Simon et al. 2008) or by fixing the ending or intermediate chords in probabilistically produced chord sequences (Kaliakatsos-Papakostas and Cambouropoulos 2014).

The architecture of the proposed system incorporates a simple statistical approach for preserving structural relations between remote harmonic parts, while at the same time diverse harmonies can be learned from data. Therefore, the merits of rule-based systems are preserved by learning and automatically employing intermediate and final cadences, leading to harmonisations that are structurally consistent. Additionally, the probabilistic nature of the incorporated algorithms allows for radically diverse harmonic idioms to be learned, while the generated harmonisations reflect the characteristics of learned idioms in terms of chord transitions and voicing layout. An additional advantage of the presented system is the fact that the output is a harmonic realisation with actual chord notes (not only chord symbols).

The presented harmonic learning system is trained independently on several harmonic aspects that are divided in two groups: chord generation and the voicing layout. Figure 1 illustrates this setting, where 'GCT generation' on the left block refers to the generation of chords symbols in the General Chord Type (GCT) representation (Cambouropoulos et al. (2014); Cambouropoulos (2015); see brief description in the next section), while the right block refers to the translation of GCT chords to actual music by assigning proper voicing layouts, converting the final output to MIDI notes. The oval blocks refer to modules that have been trained from data. The arrow leading from the 'GCT generation' to the 'GCT to MIDI pitches' block indicates the current generative process workflow of the melodic harmoniser: first, chord sequences in GCT form are produced and, afterwards, voicing layout is applied to the composed GCT sequences, providing the finalised output in MIDI format. In turn, the bass voice motion is first defined for the given GCT sequence and the given melody and, afterwards, the intermediate chord notes between bass and melody are fixed.

Both the 'GCT generation' and the 'GCT to MIDI pitches' blocks include modules that learn from data, giving the system the ability to express the characteristics of each learned idiom on several harmonic aspects. The GCT generation block incorporates three learning modules: (a) the 'Chord types' module which learns chord types by converting the pitches of the training harmonies to GCTs and organising them into chord type categories; (b) the 'Cadence constraints' module that learns and assigns cadences to user-defined positions of phrase endings (giving an essence of high-level structure); and (c) the constraint hidden Markov Model (cHMM) (Kaliakatsos-Papakostas and Cambouropoulos 2014) that learns first-order GCT chord transitions and performs probabilistic harmonisation given the aforementioned cadence constraints as well as user-defined chord constraints. The 'GCT to MIDI pitches' block includes the following learning modules: (a) the 'Bass Voice Leading' module that defines the motion of the bass in relation to the melody; (b) the 'bass-to-melody distances' that learns statistics about the distances between the bass and the melody for each idiom; and (c) the 'Chord inversions' module that learns statistics about the inversions of the learned GCT chords. The aforementioned voice-related modules contribute to the definition of the bass voice and afterwards, a simple algorithmic process, namely the 'GCT voicing layout' module, defines the chord notes between the bass and the melody.

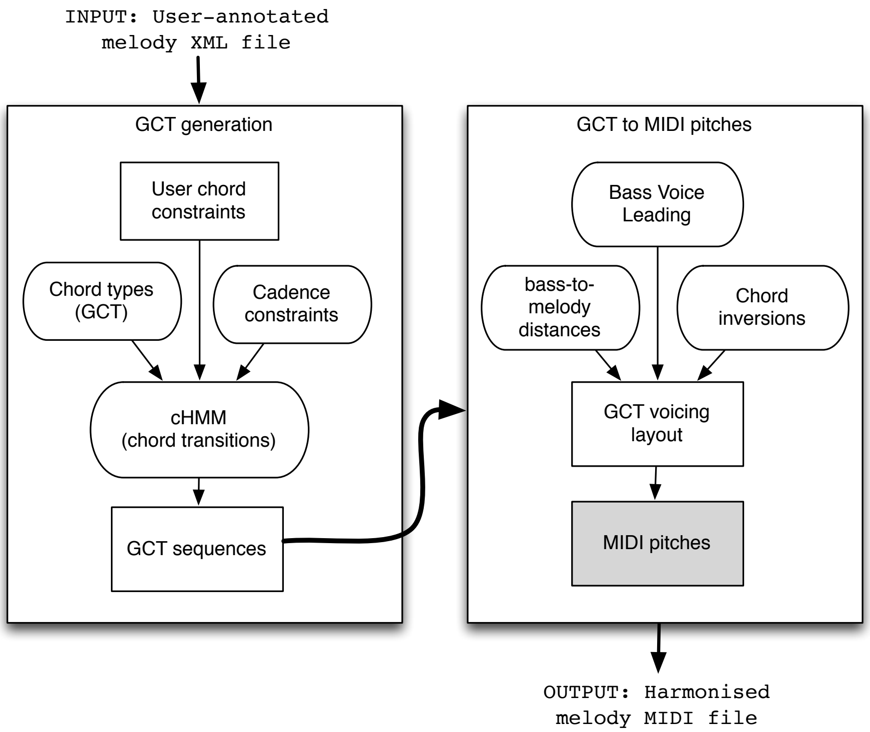


Figure 1: Overview of the Modular Melodic Harmonisation system. Oval blocks indicate modules that learn from data.

3 Chord representation and data input for training and generating

The system learns a given harmonic content given through annotated training data, while it produces new harmonisations according to guidelines provided in the melody input file. Since the processes of training and composing incorporate many diverse musical idioms, the system learns the available chord types therein (according to their root notes) based on the *General Chord Type* (GCT) (Cambouropoulos et al. 2014) representation. The training data include the notes on a level of harmonic reduction (manually annotated reductions), where only the most important harmonic notes are included, while additional 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. The chord representation scheme, the format of the training data and the user melodic input file are analysed in the following subsections.

3.1 Representation of harmony in diverse idiom with the General Chord Type encoding

The *General Chord Type* GCT provides accurate harmonic representation in the sense that it encompasses all the pitch-class-related information about chords. At the same time, for every pitch class simultaneity the GCT algorithm rearranges pitch classes so that it identifies a root pitch class and a chord 'base' (which can be considered as a basic type, e.g. major or minor) and 'extension' (which give information about chord extensions, e.g. seventh or sixth, etc.), leading to chord representations that convey musical meaning for diverse music idioms. The GCT representation has common characteristics with the stack-of-thirds and the virtual-pitch-root-finding methods for tonal music, but has differences as well (see Cambouropoulos et al. (2014)). This encoding is inspired by the standard Roman numeral chord type labelling, but is more general and flexible. A recent study (Kaliakatsos-Papakostas et al. 2015) on the GCT representation indicated that it can be used both as a means to represent harmonic chords and to describe musically meaningful relations between different harmonic labels in diverse and not necessarily tonal music idioms (Cambouropoulos et al. 2014; Kaliakatsos-Papakostas et al. 2014b; Cambouropoulos 2015).

The GCT algorithm computes, for a given multi-tone simultaneity, the 'optimal' ordering of pitches such that a maximal subset of consonant intervals appears at the 'base' of the ordering (left-hand side) in the most compact form; the rest of the notes that create dissonant intervals to one or more notes of the chord 'base' form the chord 'extension'. Since a tonal centre (key) is given, the position within the given scale is automatically calculated. Input to the algorithm is the following:

- Consonance vector: a Boolean 12-dimensional vector is employed indicating the consonance of pitch-class intervals (from 0 to 11). E.g., the vector $[1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0]$ means that the unison, minor and major third, perfect fourth and fifth, minor and major sixth intervals are consonant; dissonant intervals are the seconds, sevenths and the tritone; this specific vector is referred to in this article as the 'tonal consonance vector'.
- Pitch Scale Hierarchy: is given in the form of scale tones and a tonic. E.g., a D major scale is given as: $2, [0, 2, 4, 5, 7, 9, 11]$, or an A minor pentatonic scale as: $9, [0, 3, 5, 7, 10]$
- Input chord: list of pitch classes (MIDI pitch numbers modulo 12).

For instance, the tonic chord is labeled as $[0, [0, 4, 7]]$, where the first occurence of 0 denotes the root of the chord in relation with the scale's tonic and the base, $[0, 4, 7]$, denotes the maximally consonant setup of the included pitch classes. In relation to the tonal naming of chords, type $[0, 4, 7]$ is a major chord. Similarly the dominant seventh (inverted or not) is labeled as $[7, [0, 4, 7], [10]]$, where there is a third element, $[10]$, which is an extension, i.e. an existing pitch class that cannot be inserted in the maximally consonant set. The *compressed* GCT form will be sometimes used in this paper, where no intermediate brackets are used, e.g. $[7, [0, 4, 7], [10]]$ will be denoted as $[7\ 0\ 4\ 7\ 10]$. An example taken from Beethoven's *Sonata no. 14, op.27-2* (Figure 2) illustrates the application of the GCT algorithm for different consonance vectors. For the tonal vector, GCT encodes classical harmony in a straightforward manner. This way we have an encoding that is analogous to the standard Roman numeral encoding (Figure 2, 'tonal'). If the tonal context is changed to a chromatic scale context and all intervals are considered equally 'consonant', i.e., all entries in consonance vector are 1s, we get the second 'atonal' GCT analysis (Figure 2, 'atonal') which amounts to normal orders (not prime forms) in standard pc-set analysis. In pitch class set theory normal orders do not have 'roots' – however, they have transposition values (T0-T11) in relation to a reference pc (normally pc 0); the GCT for the 'atonal' consonance vector is equivalent to the normal orders with transposition values of pc-set theory.

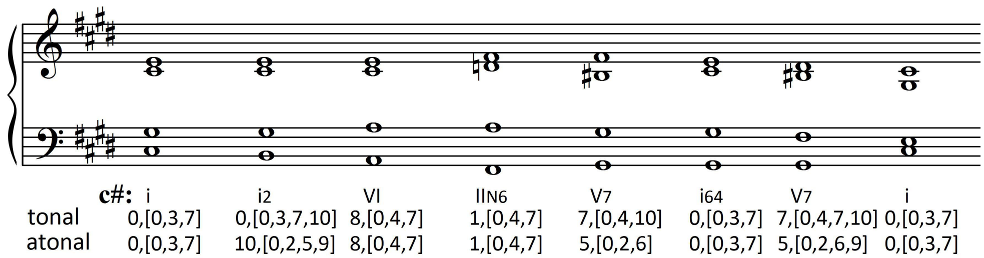


Figure 2: Beethoven, *Sonata no. 14, op.27-2* (reduction of first five measures). Top row: Roman numeral harmonic analysis; middle row: tonal GCT analysis; bottom row: atonal GCT analysis. The tonal GCT successfully encodes all chords, including the Neapolitan sixth chord (fourth chord).

An additional fundamental concern of the harmonic representation in the presented harmoniser is the grouping of chords according to their GCT representation with a methodology described in Kaliakatsos-Papakostas et al. (2015). For example, the V chord in a scale can be expressed either as $[7,\ [0,4,7]]$ or in a 'reduced' ($[7,[0,4]]$) or an 'expanded' ($[7,\ [0,4,7,10]]$) forms, that actually represent the same chord label. Each GCT group includes the GCT types that satisfy the following three criteria:

1. they have the same scale-degree root;
2. their GCT bases are subset-related; and
3. they both contain notes that either belong or not to the given scale (see Table 1 for details).

Regarding criterion 2, two bases B_1 and B_2 are considered subset-related if $B_1 \subseteq B_2$ or $B_2 \subseteq B_1$, e.g. $[0,4] \subseteq [0,4,7]$ while $[0,4] \not\subseteq [0,3,7]$. Criterion 3 is utilised to identify and group together chords that belong to secondary tonalities within the primary tonality of the piece. For instance, in a diatonic major context, while $c_1 = [0,[0,4,7]]$ and $c_2 = [0,[0,4,7,10]]$ fulfil criteria 1 and 2, according to criterion 3 they are not grouped together since c_2 includes value 10, which is mapped to the non-diatonic 10 pitch class value. In a major context $[0,\ [0,\ 4,\ 7,\ 10]]$ is secondary dominant to the IV (V/IV) and is differentiated from the I major chord.

Furthermore, each group is represented by an 'exemplar' GCT type, which is the one that is more often met in the datasets under study. Some common chord groups in the major scale Bach chorales are illustrated in Table 1. This table also includes the functional naming of each group in order to assist the comparison of the derived GCT types with the standard Roman-numeral labelling. Testing this simple algorithm on sets of both major and minor Bach chorales gives a reasonable first classification of the 'raw' GCTs. Groups of GCT chords are extracted from datasets as explained in Section 3.2 and their exemplars are used to train the system.

Table 1: Four tonal chord groups and their exemplar GCTs. The group $[0, [0, 4, 7]]$ has been separated from the group $[0, [0, 4, 7], [10]]$, due to the non-diatonic pitch class 10 of the latter.

functional name	exemplar	Group members			
tonic	$[0, [0, 4, 7]]$	$[0, [0, 4, 7]]$	$[0, [0, 4]]$	$[0, [0, 4, 7], [11]]$	
dominant	$[7, [0, 4, 7]]$	$[7, [0, 4, 7]]$	$[7, [0, 4, 7], [10]]$	$[7, [0, 4], [10]]$	$[7, [0, 4]]$
subdominant	$[5, [0, 4, 7]]$	$[5, [0, 4, 7]]$	$[5, [0, 4]]$	$[5, [0, 4, 7], [11]]$	
V / IV	$[0, [0, 4, 7], [10]]$	$[0, [0, 4, 7], [10]]$	$[0, [0, 4], [10]]$		

3.2 Training data and harmony annotations

The development of the presented melodic harmoniser incorporates statistical learning on different harmonic levels (chord transitions, cadences and voice leading) from a data pool with ‘real-world’ representations of historical traditions of music creation. By employing rich multi-level structural descriptions of harmony in different idioms, 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. Each idiom/style is internally as coherent as possible such that regularities of the specific harmonic space can be extracted; the collected idioms are as different as possible on all the examined harmonic levels. Additionally, the musical pieces are manually annotated such that structural harmonic features may be extracted at various hierarchic levels. Specifically, the following structural aspects are manually annotated: (a) harmonic reduction(s) of each musical work/excerpt so that structural harmonic/non-harmonic notes are explicitly marked; (b) local scale/key changes are determined so that harmonic concepts relating to modulations can be learnt; and (c) grouping structure is given so that cadential patterns at various hierarchic levels can be inferred.

An example of the required types of information from a music piece for training the system are illustrated in Figure 3; annotated music files include: a) the original musical data the actual musical surface and b) expert annotations that are provided as manually entered analytical information about the contents. At the lowest level of the *musical surface* (denoted as ms_0), which is the actual notes of a musical piece and the lowest level of representation that has musical significance (Jackendoff 1987), a custom text-based encoding is used. *Expert annotations* in a music piece describe musicological aspects that 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) and annotations concerning *tonality* (and tonality changes) and *grouping* information.

On the chord transitions level the system is trained according to the chord progressions on the harmonic reduction (ms_1), with chords being encoded in the General Chord Type (GCT) (Cambouropoulos et al. 2014) representation. Since the GCT requires tonality information, the GCT forms of the extracted chords are computed by using the tonality annotations. Annotations of grouping indicate the positions of cadences, where the system learns the final pairs of chords before any group ending. Even though a cadential form may incorporate more or fewer than two chords, considering the last two chords of a phrase as a cadence was decided as a rational compromise.

The indicators of the *tonality* – and the tonality changes – include accidentals in chordal form, with all the included notes indicating an octave of the utilised scale (lowest note is the tonality's fundamental), while the time instance of a tonality activation/change is defined by the indication's onset. Additionally, it has to be noted that at least one tonality indicator at the beginning of the piece is required otherwise the tonality annotations of the piece are considered absent (repetitions of the same indicator are ignored). The *grouping* part contains annotations about melodically coherent temporal regions of the music surface. At the beginning of each phrase, a group identifier is placed indicating the level of the phrase hierarchy. One note on any line indicates the lowest level groupings (e.g. phrase); two notes on the lowest two lines indicate an immediately higher-level of grouping (e.g. related phrases in a row); three notes indicate even higher level of grouping and so on. The cadences of each grouping level, i.e. the final pair of chords at the end of each grouping part, are learned separately.

Figure 3: Annotated file containing original song transcription (ms_0), time-span reduction of the original content (ms_1), as well as *tonality* and *grouping* information.

The dataset consists of over 430 manually annotated musicXML documents categorised in 7 categories and various subcategories. The separation of pieces in sets primarily focuses on genre categorisation, while subcategories are created within genres that present notable differences in their harmonic structure. The diversity in harmonic features among different sets and subcategories allows the inclusion of a wider spectrum of options, enabling the melodic harmoniser to produce harmonisations with strong references to diverse idioms. On the other hand, there is intra-idiom consistency in each subcategory of pieces, which is expressed by 'patterns' in harmonic features that are characteristic to this subcategory, in a sense that these features are often encountered in several pieces within this idiom.

The dataset comprises seven broad categories of musical idioms, further divided into sub-categories, and presented in the following list[1]:

[1]Categories 4, 5 and 6 may seem to overlap, but they are essentially different: category 4 includes harmonisations of initially monophonic folk melodies made by art music composers of European National Schools, category 5 comprises 20th-century original compositions (not based on folk songs) and category 6 contains

1. Modal harmonisation in the Middle Ages (11th – 14th centuries): includes subcategorys 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[2] 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.

For the harmonisation examples incorporated in the present paper, a subset of eight harmonic idioms was used from the dataset, presented in the following list:

1. The 15th-century Fauxbourdon style, based on parallel $^6_3$ chords.
2. The homophonic tonal harmonic idiom of J. S. Bach chorales.
3. The Kostka-Payne corpus, describing mainstream tonal harmony of the 18th/19th-centuries (Kostka and Payne 2004).
4. Edvard Grieg's 19th-century chromatic harmonic idiom, as expressed in his folk songs harmonisations op. 17 & 63.
5. The Epirus polyphonic singing style, based on the minor pentatonic scale (Lolis 2006; Kaliakatsos-Papakostas et al. 2014b).
6. Yannis Constantinidis's 20th-century modal style, as encountered in his '44 Greek miniatures for piano' (Tsougras 2010).
7. Paul Hindemith's 20th-century harmonic idiom, as expressed in his 'Six Chansons'.
8. Mainstream jazz harmony, as encountered in selected jazz standards (tonal or modal) from the *Real Book*.

original harmonisations embedded in the folk idioms.

[2]This dataset consists of the 46 excerpts that are longer than 8 measures from the workbook accompanying Kostka and Payne's theory textbook Tonal Harmony, 3rd edition (Kostka and Payne 2004) and is available in machine readable format at `http://theory.esm.rochester.edu/temperley/kp-stats/index.html`.

3.3 Melodic input

After the system is trained, it is able to harmonise a given melody. Figure 4 illustrates an instance of the input protocol for the system, which includes the melody to be harmonised and information regarding some harmonic attributes that are not inferred by the system at this stage. The input melody, in this stage, is manually annotated as to harmonic rhythm, harmonically important notes, key and phrase structure. The file that produced this figure is used as input for harmonising the example in Figure 7 (b). Initially, the user provides the positions where chords should occur (harmonic rhythm), as well as 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.

Figure 4: Example of a user melodic input file. This melodic part is from L. v. Beethoven, (b. 1-8), second movement in A♭ major of the Piano Sonata no. 8. This input file is used for the harmonised example in Figure 7 (b).

Why is tonality among the features that are specified by the user along with the melody? Although the local tonality of a melody could be automatically deduced algorithmically (Chai 2005; Kaliakatsos-Papakostas et al. 2013), manual annotation of tonality and changes has been decided for the following reasons:

1. Utilisation of non-standard (major/minor) tonalities: The collected dataset include

pieces that do not conform to the standard Western music tonalities, e.g., there are pentatonic or octatonic modes. Additionally, the user is allowed to specify any desirable tonality, which will lead the system to select the proper set of chords to harmonise the given melody.

2. Accuracy in tonality-change boundaries: Algorithms that perform melodic segmentation according to tonality (Chai 2005; Kaliakatsos-Papakostas et al. 2013) are not able to identify the exact location of tonality boundaries. For the presented melodic harmoniser, it is important that the tonality (and phrase level) change locations stay aligned with the melody segments that a human user indicates.

3. The ability to insert 'subtle' tonalities: The user is able to introduce tonality changes in places where an algorithm might not identify any change. This ability introduces additional agility and potential of variety to the system.

In the training data, tonality changes are treated differently in different idioms, while, additionally, some idioms do not include (or include very specific) modulations between certain – neighbouring in the circle-of-fifths – tonalities. Since modulations are dependent on the melody, and a user input melody might incorporate arbitrary modulations, it is clear that no learning strategy on every idiom could cover the entire spectrum of modulations that are possible for input melodies. For instance, in the idiom of modal music there are no modulations, since entire pieces are composed in a single mode. Therefore, it would be impossible to harmonise a melody that incorporates modulations using the harmony of a modal idiom, since no training paradigms would be available for such a task. For the purposes of the 'idiom independent learning' that is required for the presented system, modulations are not tackled: a cadence in the first tonality is assigned before the modulation occurs and the material after the modulation is treated as a new phrase in the new tonality.

4 Chord progressions, intermediate constraints and cadences

The core of the generative process is the production of GCT chord progressions with a probabilistic methodology that is a simple extension of the hidden Markov model (HMM) method that allows the inclusion of fixed 'anchor' chords. Harmonisation with fixed anchor chords 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 harmonised. Two types of important chords (or pairs of chords in the case of cadences) are considered: (a) intermediate or final cadences at positions where phrases end and (b) user-defined fixed chords that the system is forced to use. For the pieces used to train the system, the format of which is described in Section 3.2, annotations about phrase boundaries are also included. During training, the final pair of chords (penultimate and final chord) in each phrase is independently stored in the cadence module of the system, wherein the probabilities of intermediate and final cadences are calculated. In addition to the indicated positions of phrase endings, the user is also able to assign specific desired chords at any desired position, directly allowing the involvement of the user's creativity in the harmonisation process. Both the phrase ending positions and the user-defined chords are included in the directions provided by the user in the melody input file, as described in Section 3.3. These chords act as 'anchors' that remain fixed for the constrained HMM (cHMM) (Kaliakatsos-Papakostas and Cambouropoulos 2014) algorithm that selects *'proper'* chord sequences connecting the intermediate parts between the fixed chords, under the conditions introduced by the melodic material to be harmonised. The results presented in Kaliakatsos-Papakostas and Cambouropoulos (2014) indicate that cHMMs produce harmonisations that are potentially completely different to the ones produced by HMMs, depending on the imposed constraints.

4.1 Probabilistic generation of chord progressions with intermediate constraints

The proposed harmoniser uses the cHMM (Kaliakatsos-Papakostas and Cambouropoulos 2014) algorithm for generating chord progressions. The aim of this algorithm is to preserve the merits of probabilistic harmonisation, i.e., ability to train on different idioms and flexibility in generation, while allowing prior determination of intermediate chords (also named as *checkpoints* in the literature; see Chuan and Chew (2007)). Such constraints in the context of Markov chains (with no demands imposed by observations) are also know as 'unary' constraints (Pachet et al. 2011), however the cHMM algorithm works under the assumption sequences of states (chords) are composed given a set of observations (melody). Allowing fixed intermediate chords introduces two advantages for the presented harmoniser: (a) the preservation of higher level harmonic structure by the imposition of intermediate and final cadences and (b) the interactivity with the user by allowing any desired chord to be placed at any position. In the case of the cadences, the intermediate chords that comprise the cadence are specified by a probabilistic algorithmic process described later, that captures statistics about cadence occurrences either in intermediate phrase endings or at the end of the piece, allowing the learning of music structure on a higher hierarchical level. Direct human intervention on selecting desired chord constraints in the cHMM algorithm allows the presented harmoniser to function as a melodic harmonisation assistant that allows its user to specify a harmonic 'spinal chord' of anchor chords that are afterwards connected by chord sequences that give stylistic reference to a learned idiom.

The cHMM methodology divides the problem of finding intermediate constraints (i.e. fixed chords specified by the user or by the cadence module) into several consecutive problems of finding boundary constraints, i.e. fixed beginning and ending chords. Table 2 illustrates this process, where the intermediate chord constraints (I_j) are preserved while new chords (C_i^j) are generated, given the melody notes (m_i). The problem of assigning intermediate chord constraints is transformed into the problem of finding consecutive beginning and ending chords for each intermediate segment. In Simon et al. (2008), the HMM variation that was presented included an additional layer of probability distributions for beginning and ending chords for harmonising a part. In the cHMM methodology, used in the presented harmoniser, the probability values in the distributions for beginning and ending chords in each intermediate segment are actually binary: the chord that is selected as constraint has probability value 1, while all the others have 0.

During the cHMM training phase, an initial set of music phrases is considered which provides the system with the required statistical background, constituting the training set.

Table 2: The melody notes (m_i) need to be harmonised according to the intermediate chord constraints (I_j). The cHMM algorithm breaks the problem into two boundary constraints problems and composes the most probable sequence of chords (C_i^j) according to the observed melody, transition probabilities and given constraints.

melody notes	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
intermediate constraints	I_1				I_2			I_3
		C_1^1	C_2^1	C_3^1		C_1^2	C_2^2	

$$\underbrace{\underbrace{\qquad\qquad\qquad\qquad}_{\text{boundary constraints } 1}\underbrace{\qquad\qquad\qquad\qquad}_{\text{boundary constraints } 2}}_{\text{cHMM}}$$

For the reminder of this section, 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 notes (playing the role of 'observations' in the HMM methodology) will be denoted as $\mathcal{Y}$, while Y and y denote melody notes. There are four factors the the cHMM algorithm needs to generate a chord sequence, given a melody. Four factors are induced by the statistics from the training set.

1. The probability that each state (chord) is a beginning chord. This distribution is computed by examining the beginning chord for each phrase in the dataset and is denoted as $\pi(C_1 = c)$, $c \in \mathcal{S}$.

2. The probability that each state (chord) is an ending chord. This distribution is computed by examining the ending chord for each phrase 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, $c_{i-1} \in \mathcal{S}$. One additional 'pseudo-distribution' is included, except from the beginning and ending chords and transition probabilities learned from data.

4. A vector assigning 'pseudo-probability' values to chords that include the melody's important notes for each chord segment, denoted by $P(C_i = c_i|Y_i = y_i)$. As discussed in further detail in Section 3.3, a chord might be harmonising a phrase segment that includes more than one melody notes, while the user is able to select which among the melody notes are important. For assigning a proper chord over a melody segment, the harmoniser tries to find chords that include as many of the important notes as possible. Thereby, for each melody segment to be harmonised by a chord, each chord is assigned with a 'pseudo-probability' value according to how many of the segment's important notes it includes. Therefore, for a melody segment, chords that include more important melody notes are more probable.

The overall probability for selecting a chord in a segment of T chords is computed by

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

where

$$P_\pi = \pi(C_1 = c_1)\ P(C_1 = c_1|Y_1 = y_1), \tag{2}$$

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

$$P_\tau = \tau(C_T = c_T)\ P(C_T = c_T|Y_T = y_T). \tag{4}$$

The generated sequence of chords is statistically optimal, in a sense that it presents a maximal combination for the probabilities in all the counterparts (P_π, P_μ and P_τ), typically through the Viterbi Forney (1973) 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. However, if the beginning and/or ending chord is a constrained chord, the P_π and/or P_τ distributions respectively become 'binary', promoting only the chord that has been selected as constraint (with probability value 1). 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:

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

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

By allowing the imposition of final or intermediate chord constraints, the system is allowed to explore new harmonic paths that are suboptimal, but potentially more interesting. The relations between statistical optimality and musical interestingness is an interesting subject of research, that is left for future work.

4.2 Learning and assigning intermediate and final cadences

The limited memory according to the order of the Markov-based methods Pachet et al. (2011) does not allow them to consider longer time dependencies, a fact that is necessary for reflecting hierarchical structure of harmony. The intermediate chord constraints, as well as allowing direct user intervention in the generative process, offer the possibility to assign harmonic information in distant events, by employing intermediate and final cadences according to the phrase boundaries indicated by the user in the melodic input. Statistics for these cadences are learned during the training process (see Section 3.2), where expert annotated files including annotations for phrase endings are given as training material to the system.

Cadences are considered to be the final two chords of a phrase; during the cadence training process the two final chords in every phrase of every piece in the training data are captured. Statistics for unique cadences/pairs of chords are collected for two types of cadences:

1. *Final* cadences that are taken from the end of each piece's final phrase and
2. *Intermediate* cadences that are taken from the ending of every non-final phrase in each piece.

The final cadences collected from a set of 31 Bach chorales, which is a well-know idiom, are demonstrated in Table 3, along with the number of times they have been used. The set of final cadences collected from this set of Bach chorales reveals the specificity of cadential patterns in this idiom, including only variations of the perfect (and the use of the Tierce de Picardie for the minor). The number of different intermediate cadences is not overwhelmingly large: except for the perfect and half cadences, there are also some occurrences of the plagal and deceptive cadences along with some isolated cadential schemes that appear rarely.

Table 3: Number of occurrences of cadences induced from a set of Bach Chorales.

Final cadences			
Penultimate chord	Major scale final chord	Minor scale final chord	
	$[0\ 0\ 4\ 7]$	$[0\ 0\ 3\ 7]$	$[0\ 0\ 4\ 7]$
$[7\ 0\ 4\ 7]$	8	1	2
$[7\ 0\ 4\ 7\ 10]$	13	1	6

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 (see Section 3.3). 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. Specifically, for an intermediate or a final cadence, cadence likelihoods are taken from the probability distributions of each cadence in an idiom, i.e., how many times a cadence occurred over the total number of cadences. The appropriateness of a cadence according to the melody notes that the cadence's chords harmonise, is measured for each chord separately with the same method used in the cHMM algorithm, however, if a cadence chord lacks at least one important melody note in the segment it harmonises, then the cadence is disregarded as a whole (pair of chords). If for a given phrase ending no complete cadence (as a pair of chords) is found, then only the final chord is determined. If the utilisation of solely the final chord fails again, then no constraint is assigned for the cHMM. The motivation behind this cadence rejection mechanism was the reflection of the notion that the selected harmonisation idiom does not include a phrase closure toolset for the notes of the melody in the location that the user indicated a phrase ending – or at the end of the piece.

5 Bass voice leading and voicing layout of chords

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 (Whorley et al. 2013). This conclusion was made through the information theoretic measure cross-entropy, when the soprano, alto, tenor and bass voice were pairwise compared regarding their statistical relations. The proposed harmoniser uses a modular methodology for determining the bass voice leading presented in Makris et al. (2015b), which includes independently trained modules that function on the previously defined GCT chords that constitute the harmonisation. These modules 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 generation of chords (in GCT form) is performed by the cadence and cHMM probabilistic modules thus the selection of the proper voice layout scenarios for each GCT chord depends on the chords' inversion probabilities. After the bass voice is defined, the voicing layout of the internal chord notes is fixed.

5.1 Defining Bass Voice Motion

For constructing the bass voice leading, it is assumed that the bass voice is both a melody itself and it also depends on the piece's melody, a fact that motivates the utilisation of HMM. The primary module for defining bass motion plays the role of the hidden states under the first-order Markov assumption for bass contour (a bass motion depends on its previous one), in combination with the observations of the melody's contour (a bass motion depends on the underlying melody motion). Both the bass and the melody voice steps are represented by abstract notions that describe general quantitative information on pitch direction, also called 'direction descriptors'. In Makris et al. (2015a) several scenarios for voice contour refinement were examined, providing different levels of accuracy for describing the bass motion in different datasets. The proposed harmoniser the melody and bass note changes are divided in seven steps, as demonstrated in Table 4. The selected scenario of seven steps is based on the assumption that the perfect fourth is a small leap while the perfect fifth is a big leap.

refinement	short name	range (semitones)
steady voice	`st_v`	$x = 0$
step up	`s_up`	$1 \leqslant x \leqslant 2$
step down	`s_down`	$-2 \leqslant x \leqslant -1$
small leap up	`sl_up`	$3 \leqslant x \leqslant 5$
small leap down	`sl_down`	$-5 \leqslant x \leqslant -3$
big leap up	`bl_up`	$5 < x$
big leap down	`bl_down`	$x < -5$

Table 4: The pitch step and direction refinement scale considered for the development of the utilised bass voice leading system.

The level of refinement for representing the bass and melody voice movement give us the number of states and observations. According to the HMM methodology, the training process incorporates the extraction of statistics about the probabilities that a certain state (bass direction descriptor) follows another state, given the current observation element (melody direction descriptor), independently of the chord labels. These statistics are extracted from the training pieces of each idiom and incorporate four aspects:

1. The probability for each bass motion to begin the sequence.

2. The probability for each bass motion to end the sequence.
3. The probability that each bass motion follows another (transition probabilities).
4. The probability of a bass motion to be present, given a melody step.

The sequence of states that is generated by an HMM system, is produced according to the maximum probability described by the product of the aforementioned statistics, given a sequence of melody contour observations. The extracted probabilities for each possible next bass motion are stored in a vector of probabilities $\vec{p_m}$, which is afterwards utilised in the product of probabilities from all modules in Equation 7.

The bass voice motion provides abstract information about the motion of the bass, however, assigning actual pitches for a given set of chords requires additional information. Additionally, it might be the case that the best bass motion selected from the HMM module does not match other criteria concerning the chords that have already been selected by the cHMM, or the limits of permitted bass note pitch height. What if the best bass motion cannot be implemented for a chord, because it requires a rather improbably inversion of this chord (e.g., a V in second inversion)? What if the best bass motion drives the bass voice too high and close to the melody or too low? In order to assign a bass voice to a chord, additional information are required in the voice layout modules of the presented methodology, namely about *inversions* and the *melody-to-bass distance*. 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.

All the *inversions* of a chord are obtained by assigning each of its pitch classes as a bass note. For instance, the chord with pitch classes $[0, 4, 7]$ has three inversions, with each one having a bass note that corresponds to a different pitch class. The voicing layout module of the harmonic learning system regarding chord inversions, is trained through extracting relevant information from every (GCT) chord every piece from each music idiom. For mapping pitch-class-related inversion information directly to GCT chords, a GCT chord is considered in the form $g = [r,\ \vec{t}]$, where $\vec{t}$ is the vector describing the type of the chord, i.e. its GCT base and extension in one array. For instance, the V chord in a key is expressed as $g = [7,\ [0, 4, 7, 10]]$ in the GCT representation, where 4 denotes the major third and 7 the perfect fifth and 10 the minor seventh. In this context, the GCT type is a set of integers, $\vec{t} = [t_1,\ t_2,\ \dots\ ,\ t_n]$, where n is the number of type elements, that can be directly mapped to relative pitch classes (PCs). The statistics concerning chord inversion are expressed as the probability (p_{i}) that each type element in g is the bass note of the chord, or

$$p_{\mathrm{i}} = (v_1,\ v_2,\ \dots\ ,\ v_n),$$

where v_i, $i \in \{1, 2, \dots, n\}$, is the probability that the element t_i is the bass note. Table 5 demonstrates the extracted statistics for inversions for the most often met chords of the major mode Bach chorales. Therein it can be observed that the these chords are more often met in root position, while they are rarely played in the second inversion (fifth as bass note). Therefore, by integrating the inversion probabilities (p_{i}) within the voice layout modules as described in Equation 7, for instance the second inversion of the $[7, [0, 4, 7]]$ chord would be avoided when harmonising the style of the Bach chorales, since the probability related to its fifth being the bass note is zero.

GCT chord	relative PC	inversions
$[0, [0, 4, 7]]$	$[0, 4, 7]$	$[0.74,\ 0.23,\ 0.02]$
$[7, [0, 4, 7]]$	$[7, 11, 2]$	$[0.78,\ 0.22,\ 0.00]$
$[5, [0, 4, 7]]$	$[5, 9, 0]$	$[0.65,\ 0.34,\ 0.01]$

Table 5: Probabilities for chord inversion (p_{i}) in the three most frequently used chords in the major-mode Chorales of Bach.

An additional important aspect of voice layout concerns the absolute range of chords in the chord sequences of an idiom, i.e. the absolute difference between the bass voice and the melody. Different idioms encompass different constraints and characteristics concerning this voicing layout aspect, according to several factors, e.g., the utilised instruments' ranges. For capturing the distances between melody and bass pitch height in an idiom, interval-related information is extracted as approximate indicators about the expected pitch height of the bass voice through histograms of all melody-to-bass intervals found in the idiom's training pieces. Since exact intervals are scale-sensitive, e.g. different scales potentially produce different distributions of melody-to-bass intervals, the 'expected' bass pitch height is approximated by a normal distribution (denoted by p_{h_x}) that is adjusted to fit the distribution of the melody-to-bass intervals observed in the dataset.

For defining the pitch value of the bass in every step, the probabilities gathered from all the modules described hitherto are combined into a single value, computed as the product of all the probabilities from all the incorporated modules. To this end, for each GCT chord (C) in the sequence composed by the cHMM and cadence modules, every possible scenario of chord inversions and bass note pitch height, denoted by an index x, is generated. For each scenario (x), the product ($b_x(C)$) of all the modules discussed so far is computed, i.e. the bass motion ($p_{m_x}(C)$), the inversions ($p_{i_x}(C)$) and melody-to-bass interval $p_{h_x}(C)$:

$$b_x(C) = p_{m_x}(C)\ p_{i_x}(C)\ p_{h_x}(C). \tag{7}$$

Therefore, the best scenario (x_{best}) for the bass voice of chord C is found by: $x_{\text{best}} = \arg\max_x(b_x(C))$.

It has to be noted that the bass note motion probability ($p_{m_x}(C)$) of all examined inversions and pitch heights is obtained by the HMM module and takes a value given by the vector $\vec{p_m}$ according to the bass step it leads to. Therefore, the HMM probabilities are not utilised to compute the best sequence of all bass motions throughout the harmonisation, i.e., using the Viterbi algorithm. Contrarily, for the bass motion that is currently examined, all seven probabilities are calculated and stored in $\vec{p_m}$, while all possible pitch heights of the current chord (indexed by x) are assigned with a probability value accordingly. It should also be noted that the exact pitch height of the first bass in the first chord is calculated without information from the bass motion module ($p_{m_x}(C)$) since there is no motion in the bass before that.

An additional adjustment concerning the melody has to be made to avoid 'abnormal' for the idiom bass fluctuations in the selection of the optimal bass pitch height that are caused by potential large skips in the melody. For instance, a given melody may at some point move suddenly to very high pitches and then return to where it previously was. The effect of the melody-to-bass distribution would be to 'drag' the bass notes and make them follow the melody, producing a bass motion that sounds unnatural to most tested idioms. To this end, the melody line is 'smoothened' with a moving average of 10 positions, i.e., every pitch height in the melody is substituted by the mean value of its 10 previous pitch heights (or fewer than 10, for melody notes before the 10th).

5.2 Defining the chord notes between the bass and the melody

Obtaining the best scenario for bass voice leading determines the exact pitch value of the bass voice for each GCT chord according to the bass motion HMM, inversions of the given GCT chord and the distance between the bass voice and the melody. Depending on the number of notes in each GCT, the voicing layout, i.e. exact pitches for all chord notes, for each chord is defined. To our knowledge, no study exists that focuses on examining the position of inner voices in a generated chord. To this end, a simple statistical model is proposed that utilises a generic tree data structure to find the best combination of the intermediate voices for every chord according to some simple criteria. Our proposed methodology summarises as follow:

1. Find all the possible combinations of the intermediate notes and store them in a generic tree structure.
2. Calculate the cost for every combination and select the best.

The total cost of every combination, in turn, is based on a weighted combination three cost criteria:

- C1 Proximity to a *pitch-attractor*: The combination that best matches this criterion is the one that incorporates inner voice pitch values that are closest to a certain pitch value, named the *pitch-attractor*. The pitch-attractor value is set to a fixed ratio between the bass and the lowest melody note in the block of each chord.[3]
- C2 Evenness of *neighbouring notes distances*: Evenness in inner voices of a chord is measured by calculating the standard deviation of their pairwise distances.
- C3 Inner voice *movement distances between chords*: The inner voice movement between the previous and the current chord is calculated as the mean value of distances between the highest and the lowest inner voices. The best chord according to this criterion is the one with highest and lowest intermediate note pitches that are closest to the respective ones of the previous chord.

After thorough examination of the results in many simulations, the weight of the cost criteria are respectively: 0.5, 0.2 and 0.3. The voicing layout that is selected is the one that achieves the lowest total score in the weighted cost combination value.

[3]In the current version of the harmoniser the attractor is placed a 1/3 distance between melody and bass from the melody note. Additionally, for avoiding the 'dragging' effect of sudden melodic changes, the moving average version of the melody is used.

For example, consider that the GCT chord currently examined is $[2\ 0\ 3\ 7\ 10]$ with pitch classes $[0, 2, 5, 9]$ (D minor seventh), while the previous chord was the GCT chord $[0\ 0\ 4\ 7]$ (C major). Consider also that the MIDI pitches of the chord that comes before the aforementioned one are $[48, 55, 64]$, where the melody note is not considered, i.e. 55 and 64 are the inner notes of this chord, while for the D minor seventh the bass note value calculated by Equation 7 is 50 and the current melody note is 76. There are many possibilities for arranging the current chord's (D minor seventh) inner notes. To this end, the generic tree structure illustrated in Figure 5 is generated that represents all the voicing layout possibilities. All possible voicing layouts are taken by the tree interpretation by descending each branch from the root and they are then evaluated according to the three aforementioned criteria, the results of which are shown in Table 6.

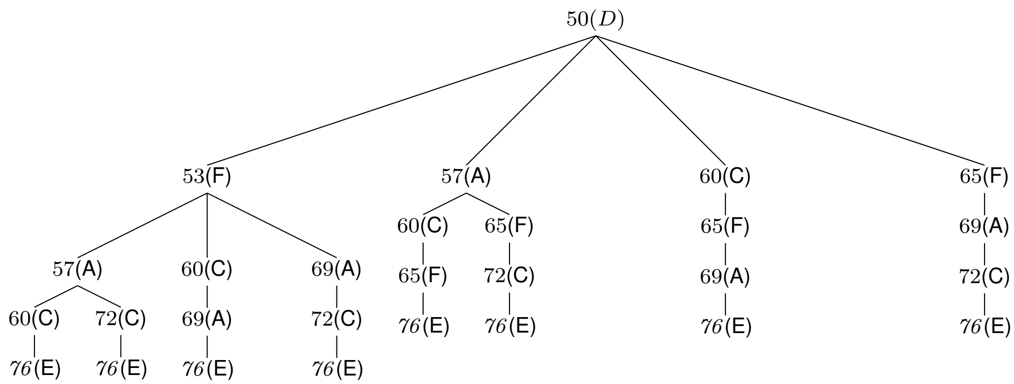


Figure 5: Tree representing all possible voicing layout scenarios for a D minor seventh with bass note 50 harmonising the melody note 76. The melody note is shown in italics since it is not actually part of the chord; it is used to demarcate the upper pitch limit of the chord's inner pitches.

Table 6: Evaluating the voicing layout scenarios for a D minor seventh with bass note 50 harmonising the melody note 76, following the chord $[48, 55, 64]$ (C major). The selected voicing layout is the one with the lowest total weighted score, shown in bold.

Voicing layout	C1 score	C2 score	C3 score	Total score
$[50, 53, 57, 60]$	6.667	0.707	6	5.275
$[50, 53, 57, 72]$	6.667	7.778	10	7.889
$[50, 53, 60, 69]$	4.667	1.414	7	4.716
$[50, 53, 69, 72]$	4.667	9.192	10	7.172
$\mathbf{[50, 57, 60, 65]}$	2.000	1.414	3	$\mathbf{2.182}$
$[50, 57, 65, 72]$	2.000	0.707	10	4.141
$[50, 60, 65, 69]$	0.000	0.707	10	3.141
$[50, 65, 69, 72]$	0.000	0.707	18	5.541

6 Experimental results

Evaluating computational or even human creativity is a difficult task, especially when the assessment of aesthetic quality is also involved. Furthermore, there is not a well-established and commonly accepted definition of creativity, as many authors approach it from different perspectives (e.g. Boden (2004); Wiggins (2006); for a comprehensive discussion see Jordanous (2013), chapter 3). The creative and structural characteristics of the system are examined through presenting examples on different harmonisation tasks as well as through statistical measures of similarities in harmonisations of melodies with different learned harmonies. The melodic harmonisation examples concern five melodies as well as different structural harmonisation attributes, e.g. intermediate phrase boundaries and user-selected chord constraints. These examples demonstrate the system's potential and indicate the integrity of harmonisations that, in some cases, reach human expert-standards with minor adjustments.

The statistical experimental process (presented in Section 6.2) examines the similarity between system-generated harmonisations of (11) different melodies and original training harmonisations. This process reveals that the harmonisations produced by the system when trained on an idiom may diverge from that idiom, depending on how its harmonic characteristics align with the structural properties and implied harmony of input melodies.

6.1 Example Harmonisations

Five diverse short melodies were chosen, three from classical music (baroque, classical and romantic periods), one from pop music and one folk song:

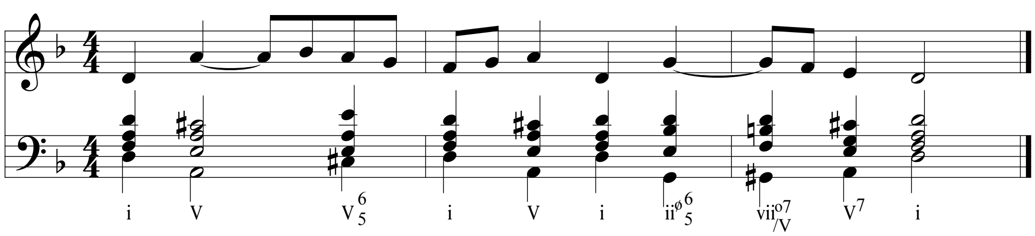
1. J. S. Bach: The fugue theme from the Well-Tempered Clavier I, fugue no. 8, transposed in D minor. The melody is a 3-bar phrase that concludes with a perfect cadence in D minor.
2. L. v. Beethoven: The melodic theme (b. 1-8) from the second movement in A♭ major of the Piano Sonata no. 8. The melody comprises two 4-bar phrases (half cadence – full cadence) that form an 8-bar period.
3. The Beatles: The first melodic phrase of the song 'Michelle', transposed in C minor. It is a 6-bar phrase, ending with a half cadence to the dominant.
4. Greek folk song: 'Tou Kitsou ē mana', taken from Yannis Constantinidis's collection '44 miniatures for piano' (no. 27). The melody is in A Dorian mode and comprises two phrases (4-bar and 7-bar) of which the second consists of two sub-phrases (3-bar and 4-bar).
5. Gabriel Fauré: The first three phrases (b. 2-21 without the repetitions) of the *Sicilienne* for cello and piano (op. 78). The melody is mainly in the Dorian mode; the first two phrases form an eight-bar period (half cadence-full cadence), while the third phrase exhibits tonal/modal mixture.

Eight different musical idioms (see Section 3.2) were used for the harmonisation of the above five melodies, but for reasons of space only a small selection of the most interesting 13 harmonisations is presented. The system produced raw midi files that were processed by humans using the Finale 2014 musical notation software[4]. The process involved the following: correction of musical notation issues and enharmonic spellings of

[4] https://www.finalemusic.com

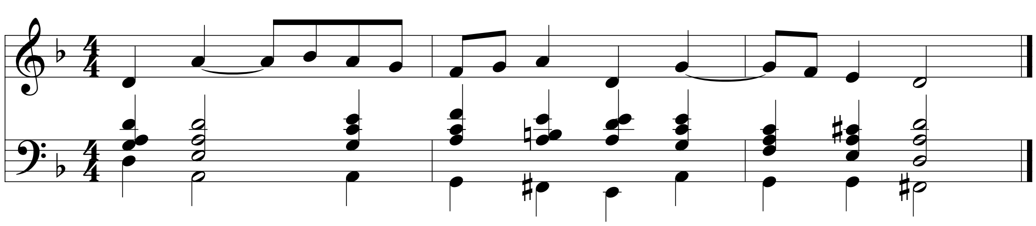
pitches, separation of the bass line in a different layer or staff, preservation of constant number of active voices in the musical texture through the use of octave doublings, manual arrangement of the inner voices for smoother voice-leading where needed, and analysis of harmonic progressions through the use of Latin Roman numeral notation in cases of tonal harmonisation. The pitch content of the chords was always kept intact, and the bass line was manually altered in very few cases (indicated by * in the scores) in order to avoid stylistic inconsistencies or achieve better voice-leading.

Three selected harmonisations of the Bach fugue are illustrated in Figure 6. The first harmonisation was based on the Kostka-Payne corpus (classical/romantic tonal harmony), which is compatible with the style of the melody, and is characterised by frequent use of the dominant and a chromatically embellished full cadence prepared by two chords with pre-dominant function: ii$^6_5$ and vii^7_o of V. The second harmonisation uses the Epirus polyphonic singing style and is almost consistently based on the D minor pentatonic scale (D, F, G, A, C) with the E of the last bar being the only exception. The chords are mildly dissonant verticalisations of the pentatonic set instead of the D minor triad, which – typically in this idiom – was avoided, and there is also a constant drone of the pitch center in the lower voice. The third harmonisation was made in the Hindemith style and exhibits free mildly dissonant chords, mostly free verticalisations of diatonic sets, except from the cadence which is tonal (V^2 - I^6). Interestingly, pitches not included in the scale of the melody are inserted for the creation of idiomatic harmony, such as B, F$\sharp$ and C$\sharp$.



(a) Bach melody harmonised in the Kostka-Payne style.

(b) Bach melody harmonised in the Epirus style.



(c) Bach melody harmonised in the Hindemith style.

Figure 6: Bach Melody harmonised in several diverse styles: (a) Kostka-Payne, (b) Epirus songs and (c) Hindemith.

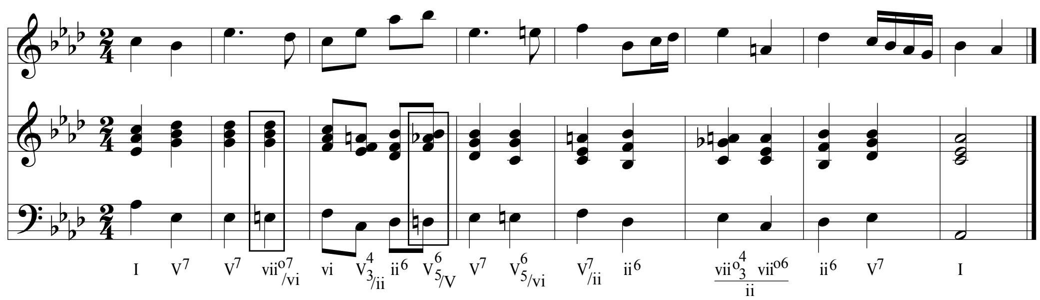
For the theme by Beethoven the three harmonisations illustrated in Figure 7 were selected. The first one (without chord constraints) was based on the Kostka-Payne idiom and is quite close to Beethoven's own style, particularly in the second phrase, which incorporates progressions in the circle of fifths and a full tonal cadence. However, the proposed harmony of the first phrase was considered static due to an insistent use of the V^7-I progression, so a second harmonisation based on the same idiom was attempted, albeit with two chord constraints in the first phrase (indicated by rectangular frames in the score). The result is substantially different, and the harmonic effect caused by the injected chords expel the tonic chord completely from the first phrase and create interesting chromatic tonicisations and a half-cadence in the phrase's end. The theme's third harmonisation used the highly chromatic Grieg idiom and rendered even more daring and interesting chromatic chords, such as the altered dominants with lowered 5^{ths} (b. 2 and 4, French-type augmented 6^{th} chords), the borrowed vii^{o7}/V with the tonic pedal note in the 3rd voice (b. 3), the tonal mixture chords $^\flat$VI and $^\flat$III (b. 5 and 6), of which the $^\flat$VI is doubly altered ($5\flat =$ B$\flat$ and $5\sharp =$ C), and the German-type augmented 6^{th} chord preparing the ii^7 chord (b. 6 and 7).

For the Beatles melodic phrase two harmonisations were selected (see Figure 8), both without any chord constraints. The first harmonisation followed the Bach chorale idiom and rendered typical diatonic or chromatic tonal progressions leading to an anticipated half cadence to the dominant. The second harmonisation was based on Yannis Constantinidis's 20th-century modal idiom, and featured almost exclusively freely used major triads with major 7^{th} and minor triads with minor 7th. In this rendering interesting parallel harmony elements are observed (A$\flat^{\text{maj7}}$-Gm^7-Fm^7-E$\flat$m), while the half cadence is avoided and substituted by a III chord with major 7^{th}. Two bass notes were manually changed (indicated by *) in order to create a complete stepwise descent from A$\flat$ to C in the bass line.

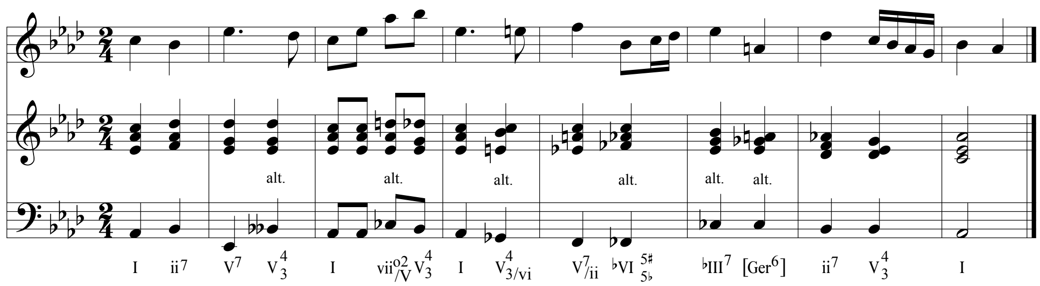
Three selected harmonisations of the Greek folk song are illustrated in Figure 9. The first was based on the Fauxbourdon medieval idiom, characterised mainly by parallel $^6_3$ chords and cadences to open 8^{th}-5^{th} sonorities. The system proposed suitable chordal content (major or minor triads, open 5^{ths} and one diminished triad as penultimate cadential chord), but the bass line had to be manually changed six times (annotated with * in the score) in order to achieve stylistic compatibility. The second harmonisation is based on Bach chorales. The result is tonal functional harmony, with diatonic and chromatic elements (tonicisations) and with tonal cadences at the end of the phrases and sub-phrases. The proposed bass line was left intact, in spite of the awkward augmented 2nd in the first bar. The last harmonisation is based on Hindemith's harmonic idiom, and is characterised by free use of chromaticism, mildly dissonant sonorities stemming from diatonic sets and more stable sonorities (major or minor triads) at the end of the phrases (a notably interesting progression is the transition from Em^7 to Gm at b. 6-7).

Finally, two harmonisations of the Sicilienne melody are illustrated in Figure 10. The first was based on the jazz harmonic idiom, characterised mainly by the free use of 7th chords and other extended/chromatic chords. The proposed harmony is a mixture of tonal and modal jazz harmony, with free chromatic or diatonic modal chords encountered during the unfolding of the melody and more tonal/functional progressions at the cadences. The second harmonisation was based on Hindemith's neotonal, mildly dissonant, nonfunctional harmony. The free chromaticism employed produced interesting enharmonic phenomena (e.g. at b. 9 and 11).

(a) Beethoven melody harmonised in the Kostka-Payne style without the user-defined chord constraints.



(b) Beethoven melody harmonised in the Kostka-Payne style with user-defined chord constraints.



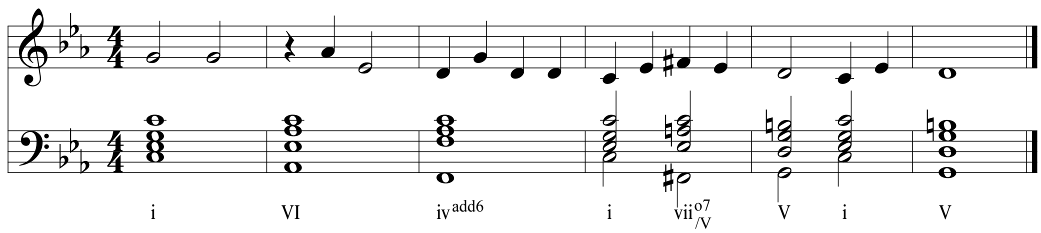
(c) Beethoven melody harmonised in the Grieg style.

Figure 7: Beethoven melody harmonised in the Kostka-Payne style (a) without and (b) with user-defined contraints and (c) in the style of Grieg.

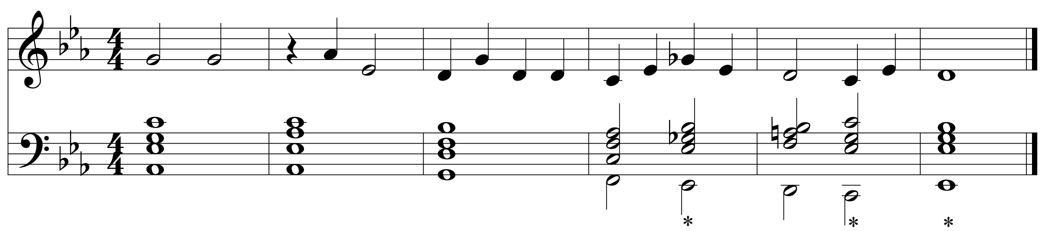
Overall, the thirteen harmonisations of the five chosen melodies produced by the system with some unobtrusive human manipulation incorporated a wide spectrum of musical idioms – with a range of over eight centuries – and demonstrated the flexibility and creative potential of the proposed harmonisation system.

6.2 Statistical similarities between original harmonies and new melodic harmonisations

The system is trained on several statistical aspect of a specific idiom and it uses the learned material in input melodies to produce novel harmonisations. How similar are the produced harmonisations in relation to the original training harmonisations of an idiom? In other words, is the system only able to mimic the training harmonisations or it is possible that 'divergent' harmonisations can be produced? This question is addressed by examining the



(a) The melody of Michelle by the Beatles harmonised in the style of Bach Chorales.



(b) The melody of Michelle by the Beatles harmonised in the style of Constantinidis.

Figure 8: The melody of Michelle by the Beatles harmonised in the style of: (a) Bach Chorales and (b) Constantinidis.

(a) Traditional melody harmonised in the style of fauxbourdon.

(b) Traditional melody harmonised in the Bach Chorale style.

(c) Traditional melody harmonised in the style of Hindemith.

Figure 9: Traditional melody harmonised in the style of: (a) fauxbourdon, (b) Bach Chorales and (c) Hindemith.

statistical similarities between original harmonies of idioms and harmonisations produced by the system for several melodies. The melodies used for producing harmonisations for

(a) The Sicilienne melody harmonised in the style of Hindemith.

(b) The Sicilienne melody harmonised in jazz style.

Figure 10: The Sicilienne melody harmonised in the style of: (a) Hindemith: and (b) jazz.

this study include the five of the ones presented previously in the examples (one major and four minor melodies), with the addition of five major mode and one minor mode melodies, to compile a total set of six major and five minor melodies. The set of major melodies includes melodies from Haydn, Mozart, Beethoven, Jobim and two traditional ones, while the selected minor melodies are by Bach, Michelle by the Beatles, Sicilienne by Fouré and two traditional melodies.

The statistical similarity of harmonies in this experimental process is based on the transitions of GCT chords. Voicing layout elements are disregarded for this study since their complex statistical interdependence makes it hard to construct a unique statistical model that can be used for statistical similarity. Instead, this study examines similarities of GCT chord transition probabilities in original pieces (used for training the system) and novel harmonisations. The examination concerns one idiom at a time, $\mathcal{I}$, where the available training harmonies (pieces with more than 4 chord transitions) are considered to form a set $\mathcal{T}_{\mathcal{I}}$, while the harmonies produced by the system for new melodies form the set $\mathcal{M}_{\mathcal{I}}$. Each harmonic piece in both sets is represented by its first-order Markov transition matrix, which represents its GCT chord transition probability distribution.

The distance between two transition probability distributions is quantified by the Hellinger distance (Gibbs and Su 2002), which is a distance metric for two distributions. Using this metric a pairwise distance matrix is constructed for both the original $\mathcal{T}_{\mathcal{I}}$ and the generated $\mathcal{M}_{\mathcal{I}}$ harmonic pieces for each idiom ($\mathcal{I}$). This matrix is mapped afterwards into a two-dimensional space using multidimensional scaling (MDS), in order to obtain a Euclidean approximation of the space of GCT chord transition distributions based on their pairwise distances. Two major and two minor-mode examples of the two-dimensional spaces produced by this process are presented in Figure 11, where the sets $\mathcal{T}_{\mathcal{I}}$ (grey $\times$s) and $\mathcal{M}_{\mathcal{I}}$ (red circles) for the Bach chorales and the Kostka-Payne sets are illustrated.

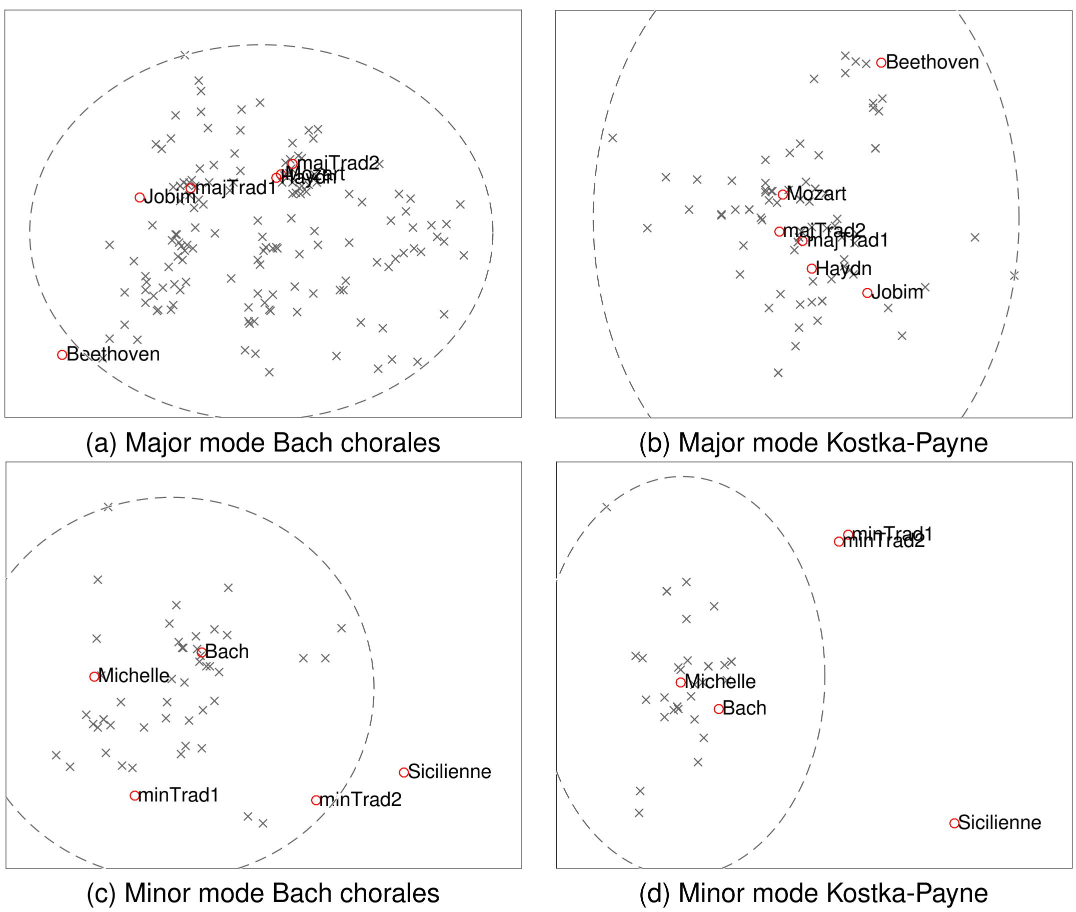


Figure 11: Examples of Bach chorales and Kostka-Payne harmonic pieces of original idiom harmonisations (illustrated with grey $\times$s) and new system-generated harmonisations (red circles) in the space produced by multidimensional scaling based on the Hellinger pairwise distances.

The original idiom harmonisation ($\mathcal{T}_{\mathcal{I}}$), as depicted in the examples in Figure 11, are considered to form a cluster. To study the relative placement of the new harmonisations in every idiom's cluster, the concept of *cluster radius* is used. Cluster radius is the maximum distance of all cluster members (harmonies in $\mathcal{T}_{\mathcal{I}}$) from the cluster centroid, which is the placed at the centre of mass of $\mathcal{T}_{\mathcal{I}}$. The radii of the clusters around their centroids are depicted by the dashed line ellipsoids in Figure 11, while the ellipticity is due to different axis scales. A harmonic sequence that is outside an idiom's radius, presents transitions in proportions that are not 'usual' (in a statistical sense) within the training idiom. The novel system-generated harmonisations ($\mathcal{M}_{\mathcal{I}}$) that are outside an idiom's cluster radius, are considered to constitute 'uncommon' new harmonisations that *explore* new harmonic areas in an idiom.

The radius for each cluster and the distances of new harmonies from the cluster's centroid are demonstrated in Table 7. One can notice that for some corpora there are more than one melodies that produce harmonisations outside the cluster's radius, e.g. in Constantinidis major and Grieg, Kostka-Payne (Figure 11 (d)), Hindemith and jazz minor. The Hindemith and jazz example harmonisations in Figure 10 of the Sicilienne melody, which are outside the respective clusters' radii, suggest that the general characteristics of the styles are locally preserved, even though the chord sequences as wholes are statistically 'divergent' from the idiom. On the other hand, all the Kostka-Payne (Figure 11 (c)) and jazz major new harmonisations are inside the cluster's radius. The music-theoretic reasons for such differences, or the perceptual impact of harmonisations outside or inside an idiom's radius are important subjects that should be addressed in future research.

Table 7: Cluster radius of intra-idiom harmonisations ($\mathcal{T}_{\mathcal{I}}$) and distances of system-generated harmonisations ($\mathcal{M}_{\mathcal{I}}$) in extra-idiom major and minor melodies from cluster centroid. Numbers is bold indicate the cases where new harmonisations are outside the radius.

	Major harmonies and melodic harmonisations						
	$\mathcal{T}_{\mathcal{I}}$ radius	Beethoven	Jobim	Haydn	Mozart	majTrad1	majTrad2
Fauxbourdon	0.4516	0.2249	0.3834	0.1341	0.2397	0.3393	**0.4851**
Bach Chorales	0.1430	**0.1541**	0.0796	0.0426	0.0462	0.0550	0.0560
Kostka-Payne	0.1398	0.0890	0.0539	0.0247	0.0192	0.0111	0.0190
Grieg	0.3350	0.2288	**0.4180**	0.1637	0.1708	0.1797	0.1728
Constantinides	0.1117	**0.2280**	**0.2913**	**0.1922**	**0.1473**	**0.2531**	**0.2542**
Jazz	0.3812	0.0449	0.1143	0.0674	0.0549	0.0852	0.0382

	Minor harmonies and melodic harmonisations					
	Intra-idiom	Bach	Michelle	Sicilienne	minTrad1	minTrad2
Fauxbourdon	0.4333	0.2852	0.1768	0.2538	**0.5492**	0.1894
Bach Chorales	0.2645	0.0626	0.1028	**0.3256**	0.1572	0.2438
Kostka-Payne	0.1670	0.0506	0.0052	**0.3413**	**0.2275**	**0.2155**
Grieg	0.3015	**0.1186**	0.0363	**0.3629**	0.0844	0.1656
Epirus	0.4193	0.0830	0.2099	0.3202	0.2586	**0.5280**
Constantinides	0.1497	0.1306	0.1148	**0.2892**	0.0451	0.1327
Hindemith	0.3111	0.1143	0.1530	**0.3850**	**0.3182**	0.2287
Jazz	0.1098	0.0541	0.0714	**0.2870**	**0.2181**	0.0882

Depending on the melody, the system may either produce harmonisations that are similar to the original training harmonies, or be forced to produce harmonisations that are less similar. This fact is important in two respects: on one hand the system is able to mimic hierarchically structured processes through a Markov-based process (using induced constraints), while on the other hand new harmonic paths can be explored. For instance, harmonising the traditional or the Sicilienne melodies with the system trained with the Kostka-Payne corpus (Figure 11 (d)), forces the system to 'explore' new harmonic areas within the idiom and generate diverse novel harmonies, in contrast to the harmonisations of the Beatles and Bach melodies. The harmonies that excess an idiom's radius, on the other hand, still reflect its learned characteristics, as indicated in the example of the Bach chorale harmonisation of the minor traditional melody 2 in Figure 9 (c), even though it is placed remotely in Figure 11 (c).

Interestingly, when the system is trained with the Bach chorales and the Kostka-Payne corpus, the relative positions of composed melodic harmonisations may be different. For instance, the harmonisations produced for the Mozart and Haydn melodies when trained with the Bach chorales (Figure 11 (a)) are very similar (one is almost placed over the other), while training the system with the Kostka-Payne corpus harmonises these melodies quite differently (Figure 11 (b)) – a fact that is possible due to the probabilistic mechanics behind the cHMM methodology. Furthermore, this is also a possible outcome in the proposed system, where even similar melodies can be harmonised in completely different ways if, for instance, different cadences are automatically selected, or, potentially, different intermediate chord constraints (or cadences) are selected by the user.

7 Concluding remarks

Melodic harmonisation with automated means is a task that requires algorithms exhibiting both emergence of creativity and preservation of structure. The first approaches for automated melodic harmonisation included methodologies that were based on human-defined rules. The strength of these approaches is that the rules they incorporate preserve the hierarchical structure of harmony. Among their shortcomings, however, is the fact that different sets of rules describe different idioms and it is impossible to come up with 'one size fits all' harmonic rules for all idioms. On the other hand, methodologies that utilise statistical learning can learn specific aspects of harmony from data, a fact that enables them to learn and create harmonies in different musical idioms. The main disadvantage of probabilistic methodologies is that they work in rather 'linear' chord-to-chord manner, disregarding higher-level structural relations between remote harmonic parts. The first contribution of the proposed melodic harmonisation system is the fact that it can learn from music data from diverse idioms, while at the same time preserve relations at distant harmonic events by assigning intermediate and final cadences at locations of phrase endings. Additionally, the system output is a complete harmonic realisation with chords being described not only as labels but as note simultaneities. To this end, different harmonic learning modules are responsible for learning and composing different aspects of harmony, namely chord types, chord transitions, cadences, bass voice movement, chord inversions and melody-to-bass note distances. Furthermore, the user can choose to import any desired chord at any location of the harmonisation, 'derailing' the system from its trained harmonic course forcing it to take creative decisions and follow alternative harmonic paths.

The creative agility of the system is obvious when used to harmonise melodies in a variety of learned idioms. Therein, the implied harmony incorporated in the melody is blended with the learned harmony employed for the harmonisation, producing interesting harmonic output. An analysis on melodic harmonisation examples, where melodies were harmonised with harmonically ‘incompatible’ learned idioms, indicated that some of these harmonisations were inside and some outside the cluster of the original idiom harmonisations. It was therefore demonstrated that the system not only reflects the characteristics of original harmonisations within an idiom, but also potentially invents new creative harmonic routes that at some extent constitute a blend of the harmonising idiom’s and the melody’s implied harmony. In other words, the system exhibits adaptivity in learning and agility in expressing learned harmonic idioms in different and potentially alien harmonic environments – as imposed by a melody’s structure. Another important aspect of the system is its ability to comply with specific user preferences in harmony, expressed as chord constraints. The user is allowed to experiment by employing desired chords in any position of the melody, forcing the system to follow potentially radically different harmonic paths in order to satisfy the user-imposed constraints. The direct involvement of the user in the creativity loop, combined with the numerous potential harmonisations using different learned idioms, make the proposed system valuable not only as an autonomous creative tool, but also as a tool that enhances the creativity of the user as a composer.

The system is developed in the context of a wider research project, where conceptual blending (Fauconnier and Turner 2003; Goguen 2006) is studied as a generative means to creating new conceptual spaces (features and relations between them) by combining the elements of two input ones. Regarding the proposed system, learned probabilistic elements of different input idioms will be transformed in logic-related feature terms, while formal computational blending processes (Schorlemmer et al. 2014; Kaliakatsos-Papakostas et al. 2014a; Cambouropoulos et al. 2015) will create new elements and relations that creatively combine and extend the input idioms by generating new probabilistic relations between them. However, the system in its current form is still a valuable tool for potential user groups. For instance, composers are able to get a ‘batch’ of creative ideas on harmonisation alternatives for a given melody within a few seconds. The system is able to provide very quickly several ideas on how a melody would be harmonised under different learned conditions, enhancing the composers’ creativity by providing many new ideas on the entire harmonisation or on parts of it. Additionally, the composers are able to keep some parts of the harmonisation fixed (as chord constraints) and search for alternatives in focused areas. Furthermore, the system can be used for educational purposes, indicating to students which harmonisation follows the most ‘usual’ harmonic paths for a given melody in diverse idioms. Students have the chance to explore creative ideas in a style-specific harmonic environment by imposing their desired chord constraints and studying the alternative harmonic routes that the system proposes in the context of a specific idiom.

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. The authors would like to thank the anonymous reviewers and the editors for their valuable ideas and comments for improving the content of this paper.

References

Allan, M. and Williams, C. K. I. (2004). Harmonising chorales by probabilistic inference. In *Advances in Neural Information Processing Systems 17*, pages 25–32. MIT Press.

Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. Psychology Press.

Boden, M. A. (2009). Computer Models of Creativity. *AI Magazine*, 30(3):23.

Borrel-Jensen, N. and Hjortgaard Danielsen, A. (2010). Computer-assisted music composition – a database-backed algorithmic composition system. B.S. Thesis, Department of Computer Science, University of Copenhagen, Copenhagen, Denmark.

Cambouropoulos, E. (2015). The Harmonic Musical Surface and Two Novel Chord Representation Schemes. In Meredith, D., editor, *Computational Music Analysis*, pages 31–56. Springer.

Cambouropoulos, E., Kaliakatsos-Papakostas, M., and Tsougras, C. (2014). 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.

Cambouropoulos, E., Kaliakatsos-Papakostas, M., and Tsougras, C. (2015). 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)*.

Chai, W. (2005). Automated analysis of musical structure. Thesis (Ph. D.)–Massachusetts Institute of Technology, School of Architecture and Planning, Program in Media Arts and Sciences, 2005.

Chuan, C.-H. and Chew, E. (2007). 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.

Conklin, D. (2002). Representation and discovery of vertical patterns in music. In *Music and artificial intelligence*, pages 32–42. Springer.

Dixon, S., Mauch, M., and Anglade, A. (2010). Probabilistic and logic-based modelling of harmony. In *Exploring Music Contents*, pages 1–19. Springer.

Donnelly, P. and Sheppard, J. (2011). Evolving four-part harmony using genetic algorithms. In *Proceedings of the 2011 International Conference on Applications of Evolutionary Computation - Volume Part II*, EvoApplications'11, pages 273–282, Berlin, Heidelberg. Springer-Verlag.

Ebcioglu, K. (1988). An expert system for harmonizing four-part chorales. *Computer Music Journal*, 12(3):43–51.

Fauconnier, G. and Turner, M. (2003). *The Way We Think: Conceptual Blending And The Mind's Hidden Complexities*. Basic Books, New York, reprint edition edition.

Forney, G.D., J. (1973). The viterbi algorithm. *Proceedings of the IEEE*, 61(3):268–278.

Gibbs, A. L. and Su, F. E. (2002). On choosing and bounding probability metrics. *International statistical review*, 70(3):419–435.

Goguen, J. (2006). Mathematical Models of Cognitive Space and Time. In Andler, D., Ogawa, Y., Okada, M., and Watanabe, S., editors, *Reasoning and Cognition*, volume 2 of *Interdisciplinary Conference Series on Reasoning Studies*. Keio University Press.

Granroth-Wilding, M. and Steedman, M. (2014). A robust parser-interpreter for jazz chord sequences. *Journal of New Music Research*, 0(0):1–20.

Hanlon, M. and Ledlie, T. (2002). Cpu bach: An automatic chorale harmonization system.

Jackendoff, R. (1987). *Consciousness and the computational mind*, volume xvi of *Explorations in cognitive science, No. 3.* The MIT Press, Cambridge, MA, US.

Jordanous, A. K. (2013). *Evaluating computational creativity: a standardised procedure for evaluating creative systems and its application*. PhD thesis, University of Sussex.

Kaliakatsos-Papakostas, M. and Cambouropoulos, E. (2014). 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.

Kaliakatsos-Papakostas, M., Cambouropoulos, E., Kühnberger, K.-U., Kutz, O., and Smaill, A. (2014a). Concept Invention and Music: Creating Novel Harmonies via Conceptual Blending. In *In Proceedings of the 9th Conference on Interdisciplinary Musicology (CIM2014)*, CIM2014.

Kaliakatsos-Papakostas, M., Katsiavalos, A., Tsougras, C., and Cambouropoulos, E. (2014b). Harmony in the polyphonic songs of epirus: Representation, statistical analysis and generation. In *4th International Workshop on Folk Music Analysis (FMA) 2014*.

Kaliakatsos-Papakostas, M., Zacharakis, A., Tsougras, C., and Cambouropoulos, E. (2015). 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.

Kaliakatsos-Papakostas, M. A., Floros, A., and Vrahatis, M. N. (2013). A clustering strategy for the key segmentation of musical audio. *Computer Music Journal*, 37(1):52–69.

Koops, H. V., Magalhães, J. P., and de Haas, W. B. (2013). A functional approach to automatic melody harmonisation. In *Proceedings of the First ACM SIGPLAN Workshop on Functional Art, Music, Modeling & Design*, FARM '13, pages 47–58, New York, NY, USA. ACM.

Kostka, S. M. and Payne, D. (2004). *Instructor's Manual to Accompany Tonal Harmony, with an Introduction to Twentieth-century Music*. McGraw-Hill.

Lolis, K. (2006). *To Epirótiko Polyphonikó Tragoúdi (Epirus Polyphonic Song)*. Ioannina.

Makris, D., Kaliakatsos-Papakostas, M., and Cambouropoulos, E. (2015a). A probabilistic approach to determining bass voice leading in melodic harmonisation. In *Mathematics and Computation in Music: Proceedings of the 5th International Conference, MCM 2015*, page Accepted for publication, London, UK. Springer, Berlin.

Makris, D., Kaliakatsos-Papakostas, M., and Cambouropoulos, E. (2015b). Probabilistic modular bass voice leading in melodic harmonisation.

Pachet, F. and Roy, P. (2001). Musical harmonization with constraints: A survey. *Constraints*, 6(1):7–19.

Pachet, F., Roy, P., and Barbieri, G. (2011). Finite-length Markov processes with constraints. In *International Joint Conference on Artificial Intelligence (IJCAI) 2011*, pages 635–642.

Paiement, J.-F., Eck, D., and Bengio, S. (2006). Probabilistic melodic harmonization. In *Proceedings of the 19th International Conference on Advances in Artificial Intelligence: Canadian Society for Computational Studies of Intelligence*, AI'06, pages 218–229, Berlin, Heidelberg. Springer-Verlag.

Pérez-Sancho, C., Rizo, D., and Inesta, J. M. (2009). Genre classification using chords and stochastic language models. *Connection science*, 21(2-3):145–159.

Phon-Amnuaisuk, S., Smaill, A., and Wiggins, G. (2006). Chorale harmonization: A view from a search control perspective. *Journal of New Music Research*, 35(4):279–305.

Phon-Amnuaisuk, S. and Wiggins, G. A. (1999). The four-part harmonisation problem: A comparison between genetic algorithms and a rule–based system. In *In proceedings of the AISB'99 symposium on musical creativity*, pages 28–34. AISB.

Raczyński, S. A., Fukayama, S., and Vincent, E. (2013). Melody harmonization with interpolated probabilistic models. *Journal of New Music Research*, 42(3):223–235.

Raphael, C. and Stoddard, J. (2004). Functional harmonic analysis using probabilistic models. *Computer Music Journal*, 28(3):45–52.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1):35–53.

Scholz, R., Vincent, E., and Bimbot, F. (2009). Robust modeling of musical chord sequences using probabilistic n-grams. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 53–56. IEEE.

Schorlemmer, M., Smaill, A., Kühnberger, K.-U., Kutz, O., Colton, S., Cambouropoulos, E., and Pease, A. (2014). Coinvent: Towards a computational concept invention theory. In *5th International Conference on Computational Creativity (ICCC) 2014*.

Simon, I., Morris, D., and Basu, S. (2008). Mysong: Automatic accompaniment generation for vocal melodies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 725–734, New York, NY, USA. ACM.

Steedman, M. (1996). The blues and the abstract truth: Music and mental models. *Mental models in cognitive science*, pages 305–318.

Suzuki, S. (2013). Four-part harmonization using probabilistic models: Comparison of models with and without chord nodes. In *Proceedings of the Sound and Music Computing Conference (SMC)*, pages 628–633, Stockholm, Sweden. Logos Verlag Berlin.

Tsougras, C. (2010). The Application of GTTM on 20th Century Modal Music: Research Based on the Analysis of Yannis Constantinidis's "44 Greek Miniatures for Piano". *Musicae Scientiae*, 14(1 suppl):157–194.

Whorley, R. P., Wiggins, G. A., Rhodes, C., and Pearce, M. T. (2013). 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.

Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7):449–458.

Wiggins, G. A., Pearce, M. T., and Müllensiefen, D. (2009). Computational modeling of music cognition and musical creativity. In *The Oxford Handbook of Computer Music*. OUP USA.

Yi, L. and Goldsmith, J. (2007). Automatic generation of four-part harmony. In Laskey, K. B., Mahoney, S. M., and Goldsmith, J., editors, *BMA*, volume 268 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Yogev, N. and Lerch, A. (2008). A system for automatic audio harmonization. In *Proceedings of the 25-th VDT International Convention*.

Zacharakis, A., Kaliakatsos-Papakostas, M., and Cambouropoulos, E. (2015). Conceptual blending in music cadences: A formal model and subjective evaluation. In *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2015)*, Malaga, Spain.

Submitted Draft - Under review

Conceptual blending of harmonic spaces for creative melodic harmonisation

Maximos Kaliakatsos-Papakostas[a], Marcelo Queiroz[b], Costas Tsougras[a], Emilios Cambouropoulos[a]

[a]School of Music Studies, Aristotle University of Thessaloniki
[b]Computer Science Department, University of São Paulo
{maxk,tsougras,emilios}@mus.auth.gr
mqz@ime.usp.br

Abstract

In computational creativity, new concepts can be invented through conceptual blending of two independent conceptual spaces. In music, conceptual blending has been primarily used for analysing relations between musical and extra-musical elements in composed music rather than generating new music. This paper presents a probabilistic melodic harmonisation assistant that employs conceptual blending to combine learned, potentially diverse, harmonic idioms and generate new harmonic spaces that can be used to harmonise melodies given by the user. The key feature of this system is the application of creative conceptual blending between the most common chord transitions (pairs of consecutive chords) of two initial harmonic idioms, while a proposed methodology integrates the blended transitions in a compound probabilistic harmonic space that preserves combined characteristics from both initial idioms along with new possible chords and transitions. This methodology enables various interesting music applications, ranging from problem solving, e.g. harmonising melodies that include key transpositions, to generative harmonic exploration, e.g. combining major-minor harmonic progressions or more extreme idiosyncratic harmonies.

1 Introduction

New concepts may be invented by traversing previously unexplored regions of a given conceptual space (exploratory creativity), transforming established concepts (transformational creativity), or by making associations between diverse conceptual spaces (combinational creativity); Boden maintains that the latter, i.e., combinational creativity, has proved to be the hardest to describe formally (Boden, 2009).

Conceptual blending is a cognitive theory developed by Fauconnier and Turner (Fauconnier and Turner, 2003) whereby elements from diverse, but structurally-related, mental spaces are combined, giving rise to new conceptual spaces: such spaces often possess new powerful interpretative properties allowing better understanding of known concepts or the emergence of altogether novel concepts. Conceptual blending is a process that allows the construction of meaning by correlating elements and structures of diverse conceptual spaces. It relates directly to Boden's notion of combinational creativity.

With regards to music, conceptual blending has been predominantly theorised as the cross-domain integration of musical and extra-musical domains such as text or image (e.g. Tsougras and Stefanou (2015); Zbikowski (2002); Zbikowski (2008); Cook (2001); Moore (2013)), and primarily discussed from a musico-analytical perspective focusing on structural and semantic integration. Blending as a phenomenon involving 'intra-musical' elements (Spitzer (2004), Antovic (2011)) is less straightforward. In principle, one of the main differences of blending theory from the theory of Conceptual Metaphor (CMT) is that it may involve mappings between incongruous spaces within a domain (e.g. conflicting tonalities in a musical composition). In this case, 'intra-musical' conceptual blending in music is often conflated with the notion of structural blending (Goguen and Harrell, 2010), and Fauconnier and Turner's theory is primarily applied to the integration of different or conflicting structural elements, such as chords, harmonic spaces, or even melodic-harmonic material from different idioms (e.g. Kaliakatsos-Papakostas et al. (2014); Ox (2014)). A more extended discussion and critical examination of conceptual blending processes in music is presented in (Stefanou and Cambouropoulos, 2015).

Different musical styles/idioms establish independent harmonic spaces that involve a network of inter-related constituent concepts such as chord, root, scale hierarchy, tonality, harmonic rhythm, harmonic progression, voice-leading, implied harmony, reduction, prolongation, and so on. Conceptual blending is facilitated when a rich background (Schorlemmer et al., 2014) of concepts is available and when these concepts are structured in such ways that creative mappings are supported. Thereby, the existence of a rich background that includes formal descriptions of diverse harmonic elements is required; the combination of concepts from different idioms injects novelty and creativity to the melodic harmonisation process.

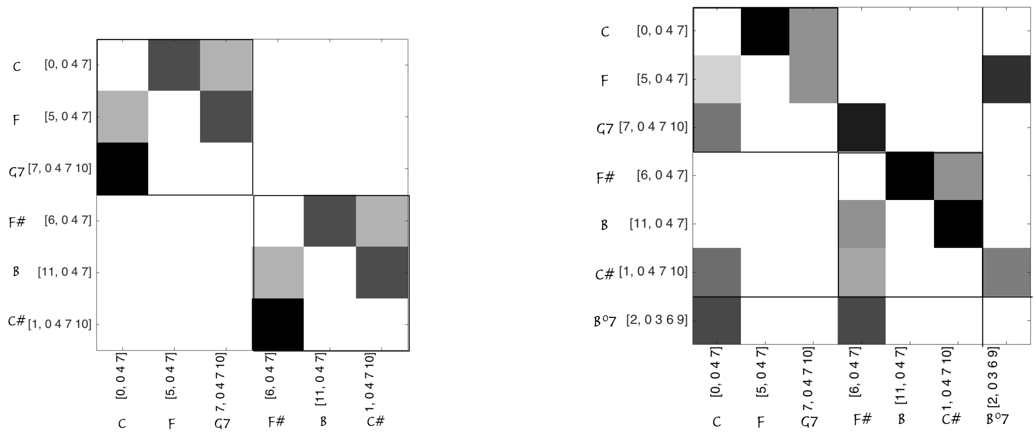
Tonal and jazz music have been effectively modelled by grammar-related structures (Rohrmeier, 2011; Koops et al., 2013; Granroth-Wilding and Steedman, 2014), however, for the purposes of blending, more musical styles need to be represented that are substantially different from the aforementioned ones. An idiom-independent representation of harmonic concepts has been proposed: from the 'primitive' chord events (see General Chord Type representation (Cambouropoulos et al., 2014)[1]) to a modular hierarchical representation of harmonic structure (Kaliakatsos-Papakostas et al., 2016b) that allows 'meaningful' blends at various hierarchic levels of harmony for practically any musical idiom. Knowledge extracted from a large dataset of more than 400 harmonically annotated pieces (manually produced harmonic reductions) from various diverse musical idioms (from medieval to 20th century styles) comprise the rich background required for interesting and creative blends. More specifically, from a set of harmonic reductions for a given idiom (e.g. Bach chorales, tango songs, jazz standards, etc.) the following structural characteristics are learned/extracted: chord types, chord transitions (probabilistic distributions), cadences (i.e. chord transitions on designated phrase endings at different hierarchic levels), and voice-leading (i.e., bass line motion in relation to melody, bass-melody distance, chord inversion). Such features from diverse idioms may be combined giving rise to new harmonic blended styles; for instance, tonal cadences may be assigned to phrase endings and modal chord transitions may be employed for filling in the rest of the phrase chords – see example in Cambouropoulos et al. (2015).

This paper focuses on the following questions: Can chord transitions per se be blended? Can two different chord transitions (e.g. cadences) from different idioms be combined to give rise to novel transitions that do not appear in any of the input harmonic spaces? Additionally, can whole chord transition matrices from different harmonic styles be amalgamated so as to generate new chord transition spaces?

Chord transition blending in the special case of cadence blending, has been explored in previous studies (Eppe et al., 2015a; Zacharakis et al., 2015). In these studies, two cadences (e.g. the tonal Perfect cadence and the modal Phrygian cadence) that share the same final tonic chord are blended giving rise to new cadences (e.g., the Tritone Substitution cadence that is commonly employed in jazz); the generated new cadences feature important characteristics from both of the input cadence spaces, namely ascending and descending leading notes to the tonic, preserving thus the closure effect of the resulting 'new' cadential formulae. In this paper, the cadence blending processes (which is based on the COINVENT conceptual blending mechanism – see Section 2 for brief description) is generalised to any two input chord transitions. This generalisation, in turn, makes possible the sophisticated creative blending of entire chord transition matrices from different idioms.

Let us attempt to illustrate the above chord transition blending processes by employing a simplistic harmonic blending example, whereby the blended spaces are merely different diatonic major tonalities. Suppose one has available (manually constructed or learned) a purely diatonic hidden Markov model on the C-major scale, with a chord (state) transition matrix complying with the first order Markov assumption and diatonic observed melodies. If a new given C major melody contains a harmonically structural F♯ note, then the Markov model reaches a dead-end as it does not know of any diatonic C major chord that can harmonise this chromatic note. If two neighbouring tonalities, however, are blended, i.e. C major and G major, then the resulting composite transition matrix contains the D major chord that leads as the dominant to the tonic in G major or as secondary dominant to the dominant in C major (see Section 3 below). For a major tonality, borrowed chords from the relative or parallel minor keys and from neighbouring tonalities can be seen as one-sided

[1]For instance, in a C major scale, the GCT representation of a C major chord is [0, 0 4 7], a G7 chord is [7, 0 4 7 10] while a B full diminished is represented as [11, 0 3 6 9].



(a) No blends (b) Topmost blends

Figure 1: Simple C major and F♯ major harmonic transition spaces with (b) no transition blends involved and (b) incorporating some of the topmost blending transitions.

blends (following Fauconnier and Turner's terminology (Fauconnier and Turner, 2003)), i.e., blends in which one primary input space remains mostly intact and specific features are imported from the secondary space.

A more extreme blend would occur between C major and F♯ major tonalities. These spaces have no common diatonic chords. Therefore the two transition matrices for these tonalities do not 'overlap' at all, and there is no way to make the transition from one space to the other. In such a case, chord transition blending may be employed to try to find new potential chord transition candidates that may allow an 'acceptable' transition between the two spaces. Let us assume that only three basic chords for each space are available, namely the tonic, subdominant and dominant seventh major chords for each space; the transition matrices for these two 'toy' spaces do not communicate (see top left and bottom right squares in Figure 1 (a)). Can the proposed chord transition methodology 'invent' new transitions that may connect the two spaces in a meaningful way?

The chord transition blending methodology is applied to all the chord transitions in the C major and F♯ major tables, i.e. each chord transition in the first matrix is blended with each chord transition in the second matrix producing a list of resulting blends. The resulting blends are ranked according to certain criteria that take into account the number of common features preserved in the blend that are shared by the input chord transitions. The features that are taken into account include common pitch classes in the first and/or second chords of the blend in relation to the two input transitions, common ascending and/or descending semitone movements in the transitions and ascending and/or descending semitone movements to the root of the final chord of each transition (see detailed description in Section 2). For instance, the transition G7 → C and C♯7 → F♯ share the same pitch classes (pcs) 5 and 11 in the first chord, have similar ascending and descending semitone movements between the two chords and contain an ascending semitone movement to the root of the final chord. Assuming that we have available a palette of basic chord types, namely, major, minor, major seventh, diminished and diminished seventh chords, a chord transition blend that ranks high is a transition in which the first chord is a diminished seventh (pcs: 2 5 8 11) and the second chord is either of C or F♯ (among other things the diminished seventh share two common pcs with each of the first chord of the input transitions). Another good blend is one where the first chord is a major seventh chord a semitone above the tonic of each space (e.g. 1 5 8 11) – this is a kind of tritone substitution transition. These invented transitions are illustrated in the new grey boxes added in the matrix of Figure 1 (b).

As seen in the above example, chord transition blending can be employed to create new transitions that preserve important features of the input transitions. When only the top ranking blends are preserved, then the system has introduced a way to connect the two input chord spaces. If more blends are selected then the composite transition matrix becomes more populated allowing more connections between the spaces. If the probabilities of the new 'invented' transitions are low, then the chord generation system creates chord sequences mostly within each of the constituent input spaces occasionally allowing passage from one to the other. If the probabilities of the new blended transitions are increased, then the whole space becomes unified and movement between most or all of the chords of both spaces is enabled. This latter strong blending between input spaces can generate new harmonic spaces that are radically different from the initial input spaces (e.g. blending two diatonic major tonalities in different keys may give rise to a composite blended space that features strong chromaticism reminiscent of music appearing centuries after diatonic tonality – see examples in Section 4).

The proposed blending paradigm seems to introduce an intelligent way to address the traditional problem of zero probability transitions in Markov models (Cleary and Teahan, 1995). Rather than assigning arbitrary non-zero 'escape' probability values (Chordia et al., 2010) or enforcing arc-consistency (Pachet et al., 2011) to allow a Markov process to cope with cases it has not seen in the training data, different transition matrices can be blended (or even a single matrix can be blended with itself) in order to introduce transitions that preserve qualities of the already existing transitions. At least for music, this seems to be a reasonable way to bypass the problem of sparse input data (e.g. learning transitions of pitch or chords or rhythmic values from a single piece rather than from a large homogeneous dataset).

In the sections below, the COINVENT blending core model will be first presented, in order to show how it is applied to chord transition blending. Then, the chord transition matrix blending methodology will be described. Finally, a number of interesting examples illustrating harmonic blending in melodic harmonisation will be given, by presenting melodies harmonised in different idioms and blends between these idioms. These results discuss different cases where harmonic blending can be useful, either as a problem solving or as a creative tool. The new possibilities offered in automated melodic harmonisation by the presented system indicate the overall usefulness of the COINVENT framework for inventing new concepts through conceptual blending. Additionally, pilot results, further investigated in a another work (Zacharakis et al., 2017), indicate that blending two harmonic spaces results in melodic harmonisations that are either perceived as belonging to a harmonic style between these two, or as belonging to a new yet intrinsically related harmonic style, fulfilling the intended purposes of blending.

2 A computational framework specialised for blending chord transitions

In computational creativity, conceptual blending has been modelled by Goguen (2006) as a generative mechanism, by describing input spaces as *algebraic specifications* and computing the blended space as their categorical *colimit*. A computational framework that extends Goguen's approach has been developed in the context of the COINVENT[2] (Concept Invention Theory) project (Schorlemmer et al., 2014). According to this framework, two *input spaces* are described as sets of features, properties and relations and after their *generic space* is computed, an *amalgamation* process (Eppe et al., 2015b; Confalonier et al., 2015) leads to the creation of several blends, which can be ranked in terms of value according to some criteria that relate to the knowledge domain of the modelled spaces.

In conceptual blending the properties of two input conceptual spaces are combined to create new spaces. The input spaces share some common structure along with differences. The intended goal of conceptual blending is to achieve a 'meaningful' combination of the non-common structural parts so that new structure emerges, giving novel properties to the generated blended space. An important aspect of the blended space is to preserve the *common* parts of the input spaces. The *generic space* is the conceptual space that keeps the common structure of the input spaces and guarantees that

[2]http://www.coinvent-project.eu

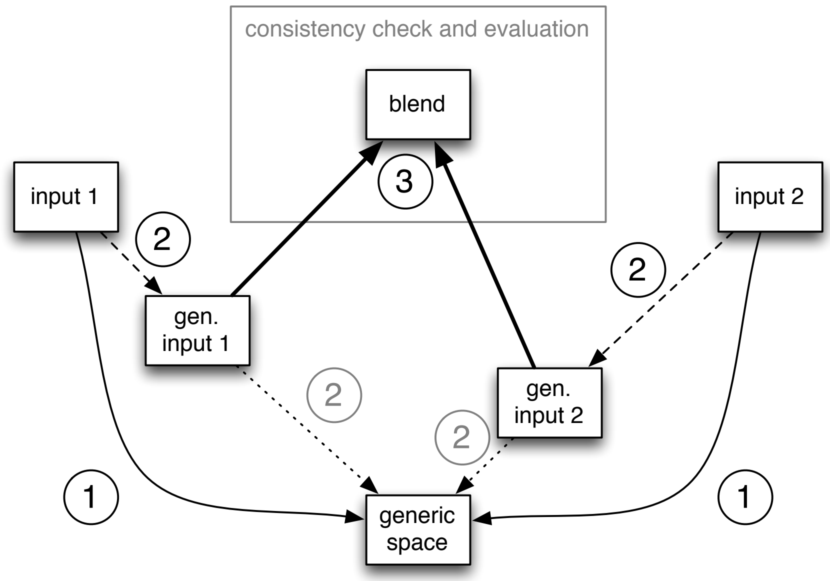


Figure 2: Conceptual blending based on amalgamation. The generic space is computed (1) and the input spaces are successively generalised (2), creating successively new potential blends (3). Some blends might be inconsistent or poorly evaluated according to blending optimality principles or domain specific criteria.

this structure also exists in the blended space. In case-based reasoning the generic space is described as 'the most specific generalisation' (Ontañón and Plaza, 2012) of the input spaces.

2.1 The COINVENT framework for conceptual blending

The COINVENT framework for generative conceptual blending is based on the notion of *amalgamation* and it is illustrated in Figure 2. An *amalgam* of two initial spaces is roughly a new space that contains parts from the initial ones (Confalonier et al., 2015). The amalgam-based workflow generalises input concepts until a generic space is found and "combines" generalised versions of the input spaces to create blends that are consistent or satisfy certain properties that relate to the knowledge domain. Figure 2 illustrates the amalgam-based COINVENT algorithmic model for conceptual blending.[3]

Amalgam-based conceptual blending has been applied to invent chord cadences (Eppe et al., 2015a; Zacharakis et al., 2015). In this setting, cadences are considered as special cases of chord transitions – pairs of successive chords, occurring before a phrase/section boundary – that are described by means of properties such as the roots or types of the chords, or specific voice motions. When blending two transitions, the amalgam-based algorithm first finds a generic space between them (point 1 in Figure 2). For instance, in the case of blending the tonal perfect cadence with the renaissance Phrygian cadences (see Eppe et al. (2015a); Zacharakis et al. (2015)) — described by the transitions I_1: G7 $\rightarrow$ C and I_2: B♭m $\rightarrow$ C5 respectively — their generic space consists of any transition that has a second chord with pitch class 0, a first chord with pitch class 5, where the first chord has a pitch class a semitone higher or lower to the second chord's root, along with other properties that might arise during blending.

After a generic space is found, 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* (Ontañón and Plaza, 2012) the properties of the inputs (point 2 in Figure 2), until the

[3] In the process of blending through amalgams, the notions of 'amalgam' and 'blend' are the same. Therefore, in the following paragraphs they are used interchangeably.

resulting unification is consistent (point 3 in Figure 2). For instance, trying to unify directly the transitions I_1: G7 → C and I_2: B♭m → C5 would yield an inconsistent amalgam, since a transition cannot both include and *not* include an upward leading note to the second chord's tonic (which is a feature of I_1 and the I_2 respectively, as discussed later). Therefore, the amalgam-based process generalises the clashing property in one of the inputs (e.g., the property describing the absence of leading note would be left empty in I_2) and tries to unify the generalised versions of the inputs again. After a number of generalisation steps are applied (point 2 in Figure 2), the resulting blend is consistent (point 3 in Figure 2). In this specific cadence blending example, one novel blend that arises from the perfect and Phrygian cadences is the Tritone Substitution progression/cadence (that is commonly used in jazz). However, it may be the case that the blend is not complete, in the sense that this process may have generated an over-generalised term. For instance, the A♭ note in the tritone substitution invention example discussed in Eppe et al. (2015a) and Zacharakis et al. (2015), is imported through completion since the C♯7 chord is required to have a perfect fifth according to their cadence formalisations, where both input chord types have a perfect fifth.

The methodology for transition blending described in the paper at hand uses an equivalent to the aforementioned methodology that combines amalgamation and completion. Chords are represented using the General Chord Type (GCT) representation (Cambouropoulos et al., 2014). The proposed methodology is adjusted for the specific harmonic ontology (with the GCT representation), using a *dictionary* of chord types that are allowed in the emerging blends. This dictionary depends on the idioms that take part in the blending process and represent a part of the "background knowledge" that these idioms incorporate. Therefore, based on the assumption that only certain chord types are allowed, the search space of possible chords in blended transitions is not overwhelmingly large, thus for the specific task of transition blending the importance of the amalgamation process is reduced and can be omitted altogether. This modification is thoroughly presented in Section 2.2.

After several blends have been computed, an evaluation process ranks them according to criteria that reflect the importance of the properties that blends inherit from the input spaces. In conceptual blending, several blending optimality principles have been proposed that are discussed in Chapter 16 of Fauconnier and Turner (2003) for rating and ranking blends. Blending quality is a necessary aspect of conceptual blending since it allows the identification of better blends among all the (potentially too) many possible ones[4]. A complete description of optimality principles is outside the scope of this paper and the reader is referred to Goguen and Harrell (2010) for applications of several such principles in the *Alloy* algorithm. The proposed methodology for rating and ranking blends in this paper is based on criteria concerning the salience of transition features within their idioms and is described in Section 2.3.

2.2 Formal description and chord transition blending

A formal ontology of transitions is required for blending according to the COINVENT framework. A chord transition (a sequence of two chords) is described as a set of properties that involve each chord independently and the chord transition as a whole (relations between the two chords). In Kaliakatsos-Papakostas et al. (2016a), an argument-based system was presented that allowed music experts to define which transition properties should be considered, through observation of blending results obtained in various harmonic setups. Using the aforementioned argument-based system and after examination of several produced outcomes, a (non-conclusive) list of nine important properties was maintained:

1. *fromPCs*: the pitch classes included in the first chord,
2. *toPCs*: the pitch classes included in the second chord,
3. *DIChas0*: Boolean value indicating whether the Directed Interval Class (DIC) vector (Cambouropoulos, 2012; Cambouropoulos et al., 2013) of the transition has 0 (i.e. that both chords have at least one common pitch class),
4. *DIChas1*: as above but for DIC value 1 (i.e., at least one ascending semitone),

[4]The amalgamation process produces many blends by following alternative generalisation paths.

5. *DIChasMinus1*: as above but for DIC value -1 (i.e., at least one descending semitone),
6. *ascSemNextRoot*: Boolean value indicating whether the first chord has a pitch class with ascending semitone relation to the pitch class of the second chord's root,
7. *descSemNextRoot*: as above but with descending semitone, and
8. *semNextRoot*: as above but with either ascending or descending semitone.
9. *5thRootRelation*: Boolean value indicating whether the first chord's root note is a fifth above the root of the second. Root notes of chords are computed with the General Chord Type (GCT) (Cambouropoulos et al., 2014) algorithm.

Table 1 demonstrates the property values (also referred to as features) for the three transitions (namely the perfect, phrygian and tritone substitution cadences) of the example discussed above. In this example, the tritone substitution cadence has been produced as a result of blending between the perfect and the phrygian cadences, with a process that is described below.

Table 1: Blending the chord transitions of the minor-mode perfect cadence (Input 1: $[7, [0\ 4\ 7], 10] \rightarrow [0, [0\ 3\ 7]]$) and the phrygian cadence (Input 2: $[10, [0\ 3\ 7]] \rightarrow [0, [0\ 3\ 7]]$). The common elements of both input spaces that are included in the generic space are depicted in boxes, while the other common elements in circles. Many blends are produced by blending these cadences; the tritone substitution blend is shown in the last column of the table as an illustrative example.

Property name	Input 1 (Perfect)	Input 2 (Phrygian)	Possible blend
fromPCs	$\{7, 11, 2, \boxed{5}\}$	$\{10, 1, \boxed{5}\}$	$\{1, 5, 8, 11\}$
toPCs	$\{\boxed{0}, \boxed{3}, \boxed{7}\}$	$\{\boxed{0}, \boxed{3}, \boxed{7}\}$	$\{0, 3, 7\}$
DIChas1	1	0	1
DIChasMinus1	0	1	1
DIChas0	1	0	0
ascSemNextRoot	1	0	1
descSemNextRoot	0	1	1
semNextRoot	1	1	1
5thRootRelation	1	0	0

In the COINVENT framework for computational conceptual blending, the role of the generic space, which includes all the common elements of the input spaces, is to reject possible blends that do not incorporate these common elements. After extensive experimentation during the development of the presented transition blending methodology, it became obvious that a richer representation of transitions that incorporates many properties potentially led to stricter generic space demands (i.e. generic spaces with more properties), allowing a smaller number of 'surprising' blends to be generated, thus reducing the creative power of the methodology. The generic space requirements are necessary for discarding blends that do not capture the important common features from the input spaces. To this end, two types of properties are distinguished: the *necessary* and the *desired* properties of transition blending. *Necessary* properties are elements incorporated in the generic space, i.e. if a necessary property is common in both inputs, then blends that don't have it are rejected. *Desired* properties are properties that characterise the input spaces and are preferred to be part of a blend, but do not belong to the generic space (i.e. they are not necessarily included in every blend). Both necessary and desired properties play an important role in rating and ranking the blends as described in Section 2.3.

In the context of the current study, among the nine properties that describe transitions, only two that concern the pitch classes of the involved chords are considered as *necessary*, namely the *fromPCs* and *toPCs*. The example in Table 1 demonstrates the role of the necessary and desired properties in transition blending. Therein, boxed items indicate the common elements in the input transitions regarded as necessary properties. For instance, all the pitch classes of the second chord as well as pitch class 5 in the first chord are present in both inputs and, therefore, are also included in all possible blends. On the other hand, the *desired* property *semNextRoot* is common in both inputs (indicated by circled numbers in Table 1); blends that do not include this property are allowed, but their rating will probably be low, depending on the salience value of this property in the inputs, to be discussed later.

So far, the discussion revolved around describing transitions with necessary and desired properties, but how are blends actually created? According to the amalgamation process, features from the input spaces should be successively generalised up to the point where no contradicting material is included (see Section 2.1 and Figure 2). This process is computationally expensive, since there are multiple generalisation paths that can be followed. Furthermore, additional musical criteria are required in order to check whether the generated blends are transitions that include 'acceptable' chord types – it is not unlikely for the algorithm to generate note clusters or trivial single-note chords that haphazardly satisfy the generic space requirements and achieve high rating value.

Considering a dictionary of acceptable chord types, it is not necessary to use amalgamation in order to explore efficiently the most prominent possible blends. By assuming that the dictionary of chord types, denoted by $\mathcal{T}$, consists of N chord types, then all the possible chords that have to be examined are 12 N – every chord type with every pitch class as a root note (i.e. all transpositions). Thereby, the 'universe' of all transitions between acceptable chords are 144 N^2. The number of acceptable types (N) is not overwhelmingly large for most musical idioms. For instance, by considering major and minor chords along with their sevenths, plus the half and full diminished chords (6 types in total), 5184 possible transitions can be generated. Therefore, producing good blends is not a matter of constructing the proper chords, but finding the ones that satisfy the necessary and desired attributes. For computing all possible blends between two input transitions, all acceptable transitions are examined regarding their compatibility with the generic space produced by the input transitions. All transitions that satisfy the generic space requirements and incorporate acceptable chord types are considered as potential blends. The algorithm for constructing the list of all possible blends is described in detail in Figure 1, while Section 2.3 analyses the process of rating and ranking all possible blends in the list.

2.3 Rating a blend

The algorithm described in Figure 3 produces a list, $\mathcal{B}$, that includes all possible acceptable transitions that are potential blends of two given input transitions (I_1 and I_2). All blends in $\mathcal{B}$ need to be rated and ranked so that *meaningful* blends are distinguished and considered with higher priority for the next steps described in Section 3. When blending two transitions taken from two different harmonic spaces, the most *meaningful* blends would expectedly include a combination of all the salient features that the input transitions encompass. The salience of a feature of a transition, however, depends on the idiom that this transition belongs to. For a set of transitions in a certain harmonic context, the more rare or characteristic a feature is, the more salient/prominent it is considered. For instance, in C major the note transition B→C ($11 \to 0$) appears in and characterises fewer chord transitions (namely G→C and Bdim→C), than say, note transition G→A ($7 \to 9$) that appears in more transitions (e.g. G→Am, C→Am, C→F, C→Dm, G→F).

To compute the salience of a feature in a transition taken from an idiom, the above mentioned 'uniqueness' of this feature needs to be quantified. To this end, let us consider the set of all transitions in an idiom, denoted by $T_{\mathcal{I}}$, where $\mathcal{I}$ is the set of indexes of all transitions in the examined idiom. Also let T_i, $i \in \mathcal{I}$ be a transition from the examined idiom. Each transition property is considered as a function of a transition, $F_p(T_i) = v_p$, returning the value of this property in a specific transition – denoted by v_p. For instance, if T_i is the perfect cadence transition (G7 $\to$ C) and $F_{\text{ascSemNextRoot}}$ is the binary function returning the *ascSemNextRoot* property value (0 or 1 for not having or having an ascending semitone to next root respectively), then the value of this property in the perfect cadence transition is obtained by $F_{\text{ascSemNextRoot}}(T_i) = 1$. We define the set of all transitions having a property p with a value v_p as

$$P_{p=v_p}(T_{\mathcal{I}}) = \{T_i, i \in \mathcal{I}; F_p(T_i) = v_p\},$$

while the cardinality (number of elements) of this set is denoted as $\mathcal{C}(P_{p=v_p}(T_{\mathcal{I}}))$. The *salience* of a property value v_p in a transition is therefore *inversely proportional* to the number of all transitions

Algorithm 1 Computation of all possible blends

Require: (**i**) two input transition, I_1 and I_2, (**ii**) a dictionary of all acceptable chord types $\mathcal{T}$
Ensure: List of all possible blends ($\mathcal{B}$) of I_1 and I_2

$\mathcal{B} \leftarrow \emptyset$ {% initialise and empty set of blends}
$g \leftarrow$ `getGenericSpace(`I_1`,`I_2`)` {% get the generic space of inputs}
$\mathcal{C} \leftarrow \emptyset$ {% initialise the set of all possible acceptable chords}
{% make the set of all possible acceptable chords}
for $t \in \mathcal{T}$ **do**
  for $r \in \{0, 1, \ldots, 11\}$ **do**
    c = `makeChordWithRootAndType(`r`,`t`)`
    $\mathcal{C}$ = `append(`$\mathcal{C}$`,`c`)`
  end for
end for
{% for all chord pairs}
for $c_1 \in \mathcal{C}$ **do**
  for $c_2 \in \mathcal{C}$ **do**
    tr = `formTransition(`c_1`,`c_2`)` {% form the transition from c_1 to c_2}
    {% check if transition satisfies generic space}
    if `satisfies(`tr`, `g`)` **then**
      $\mathcal{B}$ = `append(`$\mathcal{B}$`,`tr`)`
    end if
  end for
end for

Figure 3: Algorithm for obtaining all possible transition blends of two input transitions, given a dictionary of acceptable chord types.

in the idiom that also include this property value. Hence, the *salience*, denoted by $S_{p=v_p}(T_i)$, of a property value v_p of a transition T_i is computed as

$$S_{p=v_p}(T_i) = \frac{1}{\mathcal{C}(P_{p=v_p}(T_{\mathcal{I}}))}.$$

This salience is only defined for values v_p appearing in some transition in the idiom, from which it is immediate that the denominator above does not vanish and the feature is well-defined.

An example of applying this methodology for computing saliences is given in Table 2. The considered training idiom in this example is a set of Bach chorales in major mode, after performing GCT-based grouping (Kaliakatsos-Papakostas et al., 2015) of chords. Specifically, only the 10 most frequently used transitions of this idiom are considered, which represent each idiom as analysed in Section 3. The transitions incorporated in the example are $[7,\ 0\ 4\ 7] \rightarrow [0,\ 0\ 4\ 7]$ and $[11,\ 0\ 3\ 6] \rightarrow [0,\ 0\ 4\ 7]$ and the examined saliences concern the values of the *fromPCs* property. Both transitions include pitch class 11 as a *fromPCs* property value, but since they are the only transitions among the 10 ones representing the idiom that have it, the total salience of this feature is equally distributed among these two transitions (the 11 value of their *fromPCs* property has salience 0.5). Contrarily, the other *fromPCs* values are given smaller salience values, since they are also found in other transitions.

Table 2: Example of saliences in the respective *fromPCs* property values of two transitions in a set of major mode Bach chorales: $[7,\ 0\ 4\ 7] \rightarrow [0,\ 0\ 4\ 7]$ and $[11,\ 0\ 3\ 6] \rightarrow [0,\ 0\ 4\ 7]$. Since pitch class 11 appears as a member of the first chords only in these two transitions, the total salience of pitch class 11 in the entire idiom is equally distributed among these two.

Idiom trained on a set of major-mode Bach chorales		
example transition:	$[7,\ 0\ 4\ 7] \rightarrow [0,\ 0\ 4\ 7]$	$[11,\ 0\ 3\ 6] \rightarrow [0,\ 0\ 4\ 7]$
fromPCs property values:	$\{7,\ \mathbf{11},\ 2\}$	$\{\mathbf{11},\ 2,\ 5\}$
respective saliences within idiom:	$\{0.20,\ \mathbf{0.5},\ 0.33\}$	$\{\mathbf{0.5},\ 0.33,\ 0.25\}$

A rating value is attributed to each blend in $\mathcal{B}$ for ranking them. The rating value of a blend in $\mathcal{B}$ is computed by summing all the saliences of features that this blend inherits from the input spaces. This sum is related to the harmonic mean of the cardinalities $\mathcal{C}(P_{p=v_p}(T_{\mathcal{I}}))$ above; in fact it is precisely its reciprocal times the number of common features. Therefore, blends that incorporate a larger total of salience values inherited from the inputs are ranked as better blends, while blends that either inherit few features, or less-salient features, are ranked as worse blends.

3 Blending harmonic spaces via chord transition blending

The chord transition blending methodology described in Section 2 is integrated into the melodic harmonisation assistant presented in Kaliakatsos-Papakostas et al. (2016b). This assistant combines several probabilistic modules that learn musical structures from data, including chord transitions and cadences; chords are encoded using the GCT algorithm (Cambouropoulos et al., 2014), transitions are learned and composed with the constraint hidden Markov methodology (cHMM) (Kaliakatsos-Papakostas and Cambouropoulos, 2014), statistical models define the bass line voice leading (Makris et al., 2015) and a module fills the inner chord voices. This study attempts to employ chord transition blending in the context of the cHMM algorithm, with a view to combining creatively the independent chord transition matrices of two different harmonic idioms into a novel consistent composite harmonic space. To this end, GCT chord Markov transition tables learned from two *initial* idiom datasets are employed and most common chord transitions are indicated. Afterwards the transition blending methodology is applied on pairs of the most common transitions across the initial idioms, producing new blended transitions that connect and extend the transition possibilities of the initial idioms, generating a *compound* idiom that preserves some characteristics (in terms of transition probabilities) of the initial ones. Before transition blending is applied, a methodology for identifying potential common or similar chords of the initial idioms is employed; this enables connections between two transition tables, making musically common-sense connections between the initial idioms (see Section 3.2).

The cHMM methodology used in Kaliakatsos-Papakostas et al. (2016b) incorporates a first order Markov model, indicating the probabilities of transitions from one chord to all possible next ones. A convenient way to represent a first order Markov model is through transition matrices; this includes 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 one. Figure 4 illustrates a colour-based graphic representation of the transition matrix obtained from a collection of Bach Chorales in major mode (darker cells indicate higher probabilities). The displayed chords are actually GCT chord groups obtained by the method described in Kaliakatsos-Papakostas et al. (2015), while transitions between chords that pertain to the same GCT chord group are disregarded.

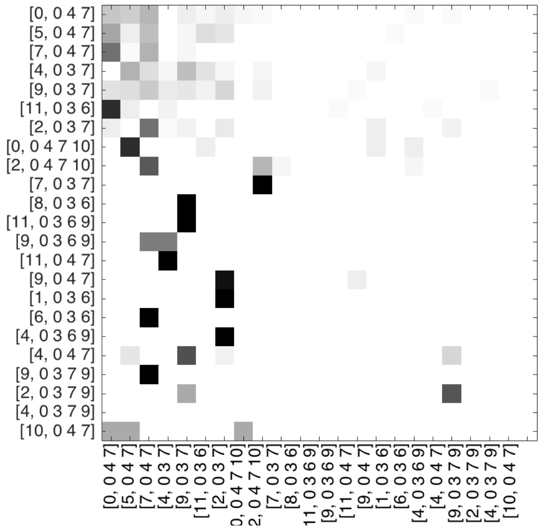


Figure 4: The first-order Markov transition matrix of chords (GCT groups) in a major Bach Chorales dataset. The numbers after the colon indicate the number of times of appearance.

3.1 A Markov transition matrix that accommodates two harmonic spaces

Aim of the proposed methodology is to construct a *musically meaningful* matrix of GCT chord transitions that includes and *extends* the respective transition matrices of two *initial* idioms. Figure 5 illustrates the general form of an *extended transition matrix* of two initial idioms. This matrix is built around the transition matrices of the initial idioms ($\mathtt{I}_1$ and $\mathtt{I}_2$), with new transitions being inserted embodying the blends that are generated by blending pairs of transitions belonging to the two initial idioms. Each initial idiom is considered to incorporate a separate set of GCT chords, even if some chords might have common names in both idioms. For instance, even though a [0, 0 4 7] chord might be found both in a Bach Chorales and a Jazz dataset, it is treated as a different chord, since it has a potentially different functional role in terms of the chords that come before or after it in each dataset. However, the identical or similar chords in the two initial idioms present “natural” harmonic connection points; transitions between such chords are constructed in a pre-blending stage, described later.

The parts of the extended matrix are the following:

1. $\mathtt{I}_1$ and $\mathtt{I}_2$: the transition matrices of the initial idioms.

2. $\mathtt{A_{i-j}}$: blended transitions that lead directly from idiom $\mathtt{I_i}$ to $\mathtt{I_j}$. For instance, a non-zero value in a cell of $\mathtt{A_{1-2}}$ enables the transition from a chord in $\mathtt{I_1}$ to a chord in $\mathtt{I_2}$.

3. $\mathtt{B_{i-X}}$: blended transitions that lead from idiom $\mathtt{I_i}$ to a new chord generated with transition blending.

4. $\mathtt{B_{X-i}}$: blended transitions that lead from a new chord generated with transition blending to idiom $\mathtt{I_i}$.

5. $\mathtt{C}$: transitions between two new chords (not considered in the current implementation and, therefore, having 0 probability value).

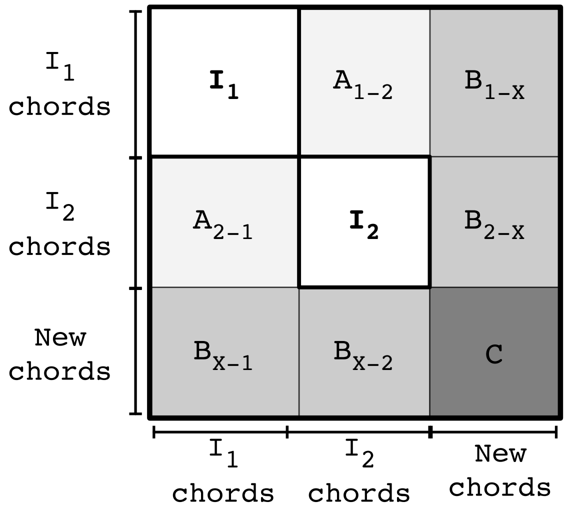


Figure 5: Graphical description of a *compound* matrix that includes transition probabilities of both initial idioms and of several new transitions generated through transition blending. These new transitions allow moving across the initial idioms, creating a new compound idiom.

The proposed process for constructing the compound matrix, intuitively generates new transitions by blending the most common transitions in the initial ($\mathtt{I_i}, i = 1, 2$) Markov matrices. It is not straightforward how a blended transition can be inserted in the extended matrix, since the extended matrix is a means to *interconnect and relate* chords between $\mathtt{I_i}$ and $\mathtt{I_j}$. The idea behind the proposed methodology is that blended transitions should allow moving from chords in $\mathtt{I_i}$ to ones in $\mathtt{I_j}$ and vice-versa. However, transition blending can potentially invent transitions that include 'new' chords that are not included in the chord sets of both initial idioms. In the case where transitions that include new chords are generated, for the proposed methodology we ensure that every transition should have at least one chord that departs from or leads to $\mathtt{I_i}, i = 1, 2$ and at most one new chord. Therefore, in this study we assume that blended transitions can include only one new 'intermediate' chord for moving from $\mathtt{I_i}$ to $\mathtt{I_j}$, while blends that include chords that are both new in both idioms are discarded. Additionally, we need to *ensure* that if a new chord is used, it should be preceded by a chord in $\mathtt{I_i}$ and be followed by a chord in $\mathtt{I_j}$. If this requirement is not met, the new chord would be either a terminal or a beginning chord, constituting a 'dead-end' or 'unreachable' chord.

By analysing the graphical representation of an extended matrix as depicted in Figure 5 the following facts need to be highlighted:

1. By using transitions in $\mathtt{I_i}$, only chords of the i-th idiom are used. When using the transition probabilities in $\mathtt{I_i}$, the resulting harmonisations preserve the character of idiom i.

2. Transitions in $\mathtt{A_{i-j}}$ enable direct jumps from chords of the i-th to chords of the j-th idiom. If a blended transition happens to be in $\mathtt{A_{i-j}}$ there is no need for further considerations – such transitions can be directly included in the extended matrix.

3. Transitions in $\texttt{B}_{\texttt{i}-\texttt{X}}$ go from a chord of idiom i to a new (in both initial idioms) chord created with transition blending. Similarly, transitions in $\texttt{B}_{\texttt{X}-\texttt{j}}$ 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 required: $c_i \rightarrow c_x$ followed by a transition $c_x \rightarrow c_j$, where c_i is in idiom i and c_j is in idiom j. A chain of two consecutive transitions with one intermediate external chord from chords of i to chords of j will be denoted as $\texttt{B}_{\texttt{i}-\texttt{X}-\texttt{j}}$.

4. Transitions in $\texttt{C}$ incorporate pairs of chords that are new to both the i-th and the j-th idioms. Having two external chords, transition blends in $\texttt{C}$ are disregarded in the present work and, therefore, all probabilities in $\texttt{C}$ are set to 0.

3.2 Connecting transition tables via common chords

Two harmonic spaces may share common chords, or chords that belong to the same GCT chord group. According to the methodology presented in Kaliakatsos-Papakostas et al. (2015), two chords belong to the same group if they (i) have the same root; (ii) have subset-related chord types; and (iii) both include pitch classes that are diatonic to the scale of the idiom. For instance, in a C major key, the chords [0, 0 4], [0, 0 4 7] and [0, 0 4 7 11] belong to the same major tonic group, while [0, 0 4 7 10] belongs to another since the pitch class value 10 is not diatonic to the major scale (this chord is a secondary dominant seventh to F major). For the remaining of this section, the term 'similar chords' will be used for describing chords that belong to the same GCT group.

The first step for generating the compound version of two transition matrices does not include blended transitions, but transitions that are composed of identical or similar chords between the two initial spaces – formulating an initial set of $\texttt{A}_{1-2}$ and $\texttt{A}_{2-1}$ transitions. These transitions allow moving between the two initial spaces by using common or similar chords as harmonic connection points. To this end, all possible transitions of such chords (i.e. all preceding and next chords) in one input idiom $\texttt{I}_{\texttt{i}}$, are also considered as possible transitions of this chord in the other input idiom $\texttt{I}_{\texttt{j}}$, "activating" the respective transitions in $\texttt{A}_{1-2}$ and $\texttt{A}_{2-1}$.

An example of this process is illustrated in Figure 6. In this example, chords $\texttt{X}_1$ in $\texttt{I}_1$ and $\texttt{X}_2$ in $\texttt{I}_2$ are similar. The previous and next chords of $\texttt{X}_{\texttt{i}}$ are shown as probability values in the vertical and horizontal stripe areas respectively in $\texttt{I}_{\texttt{i}}$, with solid striped and dotted patterns for $i = 1$ and $i = 2$ respectively. Since $\texttt{X}_1$ and $\texttt{X}_2$ are similar in both spaces, the following transitions are activated:

1. All chords leading to $\texttt{X}_1$ in $\texttt{I}_1$ (vertical solid striped pattern) should be also leading to $\texttt{X}_2$. Therefore, these probability values are copied to the faded vertical striped pattern area in $\texttt{A}_{1-2}$.

2. All chords departing from $\texttt{X}_1$ in $\texttt{I}_1$ (horizontal solid striped pattern) should be also departing from $\texttt{X}_2$. Therefore, these probability values are copied to the faded horizontal striped pattern area in $\texttt{A}_{2-1}$.

3. All chords leading to $\texttt{X}_2$ in $\texttt{I}_2$ (vertical solid dotted pattern) should be also leading to $\texttt{X}_1$. Therefore, these probability values are copied to the faded vertical dotted pattern area in $\texttt{A}_{2-1}$.

4. All chords departing from $\texttt{X}_2$ in $\texttt{I}_2$ (horizontal solid dotted pattern) should be also departing from $\texttt{X}_1$. Therefore, these probability values are copied to the faded horizontal dotted pattern area in $\texttt{A}_{1-2}$.

It should be noted that in this example, $\texttt{X}_1$ and $\texttt{X}_2$ are considered similar both in $\texttt{I}_1$ and $\texttt{I}_2$. In the case where two chords are similar only in the context of one initial space, then only the steps related to this space are followed, i.e. steps 1-2 or steps 3-4 in the enumeration given above. An example of such similarity asymmetry is the case of chords G7 and G, which are similar (belong to the same GCT group) in a C major space, while they are not in G major (since the flat seventh of G7 is not diatonic).

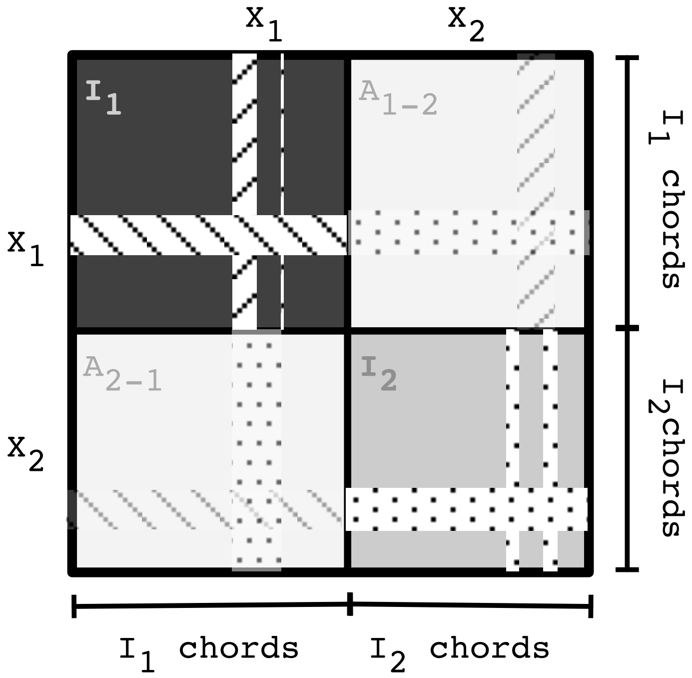


Figure 6: Graphical representation of shared transitions between similar chords ($\mathtt{X}_1$ and $\mathtt{X}_2$) in both initial spaces ($\mathtt{I}_1$ and $\mathtt{I}_2$ respectively). Transitions departing from or arriving to the similar chords within the initial spaces (solid colours in the $\mathtt{I}_1$ and $\mathtt{I}_2$ blocks) are 'mutually exchanged' in the $\mathtt{A}_{1-2}$ and $\mathtt{A}_{2-1}$ areas (faded colours) of the extended matrix, before transition blends are imported.

3.3 Identifying meaningful transition blending candidates

In order to reduce the number of of applications of chord transition blending, the 10 most common transitions in $\mathtt{I}_1$ and $\mathtt{I}_2$ are gathered in two chord transition sets that represent the respective initial idioms. Every transition of idiom 1 is blended with the ones of idiom 2, producing 100 different potential applications of blending. Some applications of blending, however, may potentially subsume others, in a sense that some transition blending pairs may incorporate harmonic characteristics that have already been examined in other pairs. For the current study, meaningful transition blends are considered the ones that incorporate *maximal* subsets of features from the generic spaces in regards to the *subsumption* relation, as explained in the next paragraphs.

Each pair of input chord transitions (x_1, x_2) defines a generic space ($\mathcal{G}_{(x_1,x_2)}$) and a set of all possible generated blended transitions ($\mathcal{B}_{(x_1,x_2)}$); the generic space represents the common properties of the two input transitions, as described in Section 2.2, and by extension it defines the set of all possible blended transitions that fulfill its requirements, generated by Algorithm 1 and later ranked as per Section 2.3. It should be reminded that in the proposed transition blending methodology, only the properties related to pitch classes are considered in the generic space, namely the *fromPCs* and *toPCs* properties.

A generic space ψ_1 is said to *subsume* a generic space ψ_2 (or ψ_2 is subsumed by ψ_1), denoted as $\psi_1 \sqsubseteq \psi_2$, if ψ_1 is more general than or equal to ψ_2 (or equivalently ψ_2 is more specific than or equal to ψ_1) in the sense that ψ_1 defines a larger set of possible generated blended transitions. Using the above notation, $\mathcal{G}_{(x_1,x_2)} \sqsubseteq \mathcal{G}_{(y_1,y_2)}$ is equivalent to $\mathcal{B}_{(x_1,x_2)} \supseteq \mathcal{B}_{(y_1,y_2)}$, which means that the pair of transitions (y_1, y_2) gives rise to a smaller set of blended transitions $\mathcal{B}_{(y_1,y_2)}$ due to a more constrained generic space (i.e. the generic space $\mathcal{G}_{(y_1,y_2)}$ is more specific than $\mathcal{G}_{(x_1,x_2)}$).

The subsumption relation between generic spaces defines a *partial order relation*, that is, in the set of all possible generic spaces the subsumption relation satisfies, for all ψ_i, ψ_j, the following properties:

1. reflexivity: $\psi_i \sqsubseteq \psi_i$;
2. antisymmetry: if $\psi_i \sqsubseteq \psi_j$ and $\psi_i \sqsupseteq \psi_j$ then $\psi_i = \psi_j$; and
3. transitivity: if $\psi_i \sqsubseteq \psi_j$ and $\psi_j \sqsubseteq \psi_k$ then $\psi_i \sqsubseteq \psi_k$.

Therefore, in every finite subset of generic spaces there is at least one *maximal element*[5] ψ_M, for which no other generic space ψ_i is more specific than ψ_M. In other words, in any set of pair of transitions there are always (one or more) pairs that define maximal (i.e. most specific) generic spaces, while other non-maximal pairs are characterized by weaker conditions (i.e. with less strict generic spaces) that produce larger blending sets containing those produced by the most specific generic spaces.

The subsumption relation is then utilised to discard blended quadruples (i.e. pairs of transitions formed by 4 chords) that are 'overshadowed' by others that incorporate a larger number of comparable common properties between the inputs. Formally, by considering the set $\mathcal{G} = \{\mathcal{G}_i\}$ of generic spaces from all blending quadruples that are generated for two initial idioms, the blends that are retained are the ones that correspond to maximal generic spaces in $\mathcal{G}$. Since each blended transition also has a rating value, this set can be further reduced by applying a threshold on rating value or on the number of desired outputs. In the next sections a maximum of 100 blends with highest rating values will be retained for further processing.

Let us consider an example where idiom I_1 is a simple C major space that includes only three transitions: C $\rightarrow$ F, F $\rightarrow$ G7 and G7 $\rightarrow$ C; let I_2 be a simple F major space with only three transitions: F $\rightarrow$ B♭, B♭ $\rightarrow$ C7 and C7 $\rightarrow$ F. Table 3 shows all transition input pairs between these two spaces and their respective generic space properties. Since these three-chord C major and F major spaces have many pitch classes and many chords in common, their extended transition matrix include blended transitions that reflect their strong relation. According to this approach, meaningful blends are considered to be the ones that belong to blending quadruples with maximal (i.e. most specific) generic spaces. In the example under discussion, the maximal generic spaces are the following:

1. $\mathcal{G}_1$, $\mathcal{G}_2$, $\mathcal{G}_6$, $\mathcal{G}_9 \sqsubseteq \mathcal{G}_3$;
2. $\mathcal{G}_1$, $\mathcal{G}_7 \sqsubseteq \mathcal{G}_4$; and
3. $\mathcal{G}_5$, $\mathcal{G}_7 \sqsubseteq \mathcal{G}_8$.

Therefore, only the blends corresponding to the blending quadruples with generic spaces $\mathcal{G}_3$, $\mathcal{G}_4$ and $\mathcal{G}_8$ are considered; the others are discarded[6].

In the case where two blending quadruples include identical generic spaces with different input transitions, the blends of both quadruples are retained. Even though, according to the algorithm in Figure 3, the blends in quadruples with identical generic spaces are the same, they are evaluated differently, since the input transitions that produced them are different. Therefore, these quadruples include blends that are ranked differently, leading to different selections of topmost blends in the subsequent steps.

3.4 Assigning probabilities and embedding transition blends in the extended idiom matrix

For each quadruple that passes the generic space subsumption filtering stage, the topmost 100 blends are kept while the rest are discarded. For each of these blends, a probability value is calculated, that will be used in subsequent selection steps described later. The proposed approach for assigning

[5] ψ_M is maximal with respect to $\{\psi_1, \psi_2, \ldots, \psi_n\}$ when there is no element ψ_i with $\psi_M \sqsubseteq \psi_i$, or equivalently, when for each i it either holds that $\psi_i \sqsubseteq \psi_M$ or ($\psi_M \not\sqsubseteq \psi_i$ and $\psi_i \not\sqsubseteq \psi_M$).

[6] For computational efficiency, in the implemented system the quadruple rejection/acceptance step precedes the blending step – therefore the blends of the discarded quadruples are actually never computed.

Table 3: Input spaces and their generic spaces in combining C major with F major three-chord spaces.

	Inputs		Generic space elements	
Index	Input 1	Input 2	*fromPCs*	*toPCs*
1:	[0, 0 4 7] $\rightarrow$ [5, 0 4 7]	[5, 0 4 7] $\rightarrow$ [10, 0 4 7]	{0}	{5}
2:	[0, 0 4 7] $\rightarrow$ [5, 0 4 7]	[10, 0 4 7] $\rightarrow$ [0, 0 4 7 10]	*ANY*	{0}
3:	[0, 0 4 7] $\rightarrow$ [5, 0 4 7]	[0, 0 4 7 10] $\rightarrow$ [5, 0 4 7]	$\{\mathbf{0},\mathbf{4},\mathbf{7}\}$	$\{\mathbf{5},\mathbf{0},\mathbf{9}\}$
4:	[5, 0 4 7] $\rightarrow$ [7, 0 4 7 10]	[5, 0 4 7] $\rightarrow$ [10, 0 4 7]	$\{\mathbf{5},\mathbf{9},\mathbf{0}\}$	$\{\mathbf{2},\mathbf{5}\}$
5:	[5, 0 4 7] $\rightarrow$ [7, 0 4 7 10]	[10, 0 4 7] $\rightarrow$ [0, 0 4 7 10]	{5}	{7}
6:	[5, 0 4 7] $\rightarrow$ [7, 0 4 7 10]	[0, 0 4 7 10] $\rightarrow$ [5, 0 4 7]	{0}	{5}
7:	[7, 0 4 7 10] $\rightarrow$ [0, 0 4 7]	[5, 0 4 7] $\rightarrow$ [10, 0 4 7]	{5}	*ANY*
8:	[7, 0 4 7 10] $\rightarrow$ [0, 0 4 7]	[10, 0 4 7] $\rightarrow$ [0, 0 4 7 10]	$\{\mathbf{2},\mathbf{5}\}$	$\{\mathbf{0},\mathbf{4},\mathbf{7}\}$
9:	[7, 0 4 7 10] $\rightarrow$ [0, 0 4 7]	[0, 0 4 7 10] $\rightarrow$ [5, 0 4 7]	{7}	{0}

probability values to the blends of a quadruple is intended to reflect (a) the probability values of the input transitions that produced these blends and (b) the ranking placement of each blend in the blending quadruple. Specifically, if the probability value (in the initial transition matrix of the idiom) of the inputs that produced a blend is p_{I_1} and p_{I_2}, then the probability of a blend, p_b, is computed as:

$$p_b = \frac{p_{I_1} + p_{I_2}}{2} \, \frac{\text{rate}(b)}{\text{rate}_{\text{max}}},$$

where rate(b) is the rating value of the blend and rate_{max} is the maximum rating value in the examined blending quadruple. In other words, the probability assigned to a blend is the mean probability of the inputs that produced it, scaled by a factor that indicates the rating of this blend in comparison to the best-rated blend that these inputs have produced – the better the rate of the blend, the closer its probability value to the mean value of probabilities of the inputs.

Among the blending quadruples that are preserved, a number of their best blends is stored for further investigation, creating a pool of best blends. Based on trials, a large number of the best blends (i.e. 100) from each blending quadruple should be inserted in the pool of best blends, so that several scenarios for connecting the initial spaces can be created, since a greater number of blends in the pool of best blends introduces a larger number of possible commuting paths in $\mathtt{A_{i-j}}$ or in $\mathtt{B_{i-X-j}}$.

Blended transitions in the pool of best blends are, then, categorised according to whether they belong to category $\mathtt{A_{i-j}}$, $\mathtt{B_{i-X}}$ or $\mathtt{B_{X-i}}$. Blends that belong to category $\mathtt{A_{i-j}}$ can be directly embedded in the extended transition matrix. However, blends that belong to either $\mathtt{B_{i-X}}$ or $\mathtt{B_{X-i}}$ may potentially constitute terminal or beginning transitions respectively, as discussed in Section 3.1. Therefore, blends in $\mathtt{B_{i-X}}$ or $\mathtt{B_{X-i}}$ are matched in $\mathtt{B_{i-X-j}}$ chains/pairs and considered as integrated elements. The rating value assigned to every chain of blended transitions is the mean of ratings of each blend in the chain.

For allowing different intensities of blending in the harmonisations that the system produces, there are also two parameters, namely *rating-based selection* (RBS) and *probability intensity multiplier* (PIM), that define the number of blends to be embedded in the extended matrix and the relative value of probabilities of transitions outside the initial harmonic spaces ($\mathtt{I_1}$ and $\mathtt{I_2}$). RBS is in the range $[0, 1]$ and defines the percentage of top blends or transition chains that are imported in the extended matrix. For instance a RBS value of 0.5 imports 50% of the most highly rated blends, while a value of 0 generates an extended matrix that includes only the initial spaces and the pre-blending common/similar connections (see Section 3.2). PIM is in the $[0, 1]$ range and is used as a multiplier of all probabilities outside the $\mathtt{I_1}$ and $\mathtt{I_2}$ according to the following formula:

$$p_{\text{new}} = (0.1 + 20PIM)\ (p_{\text{old}}^{(1-\frac{PIM}{2})}),$$

where p_{new} is a new value (potentially greater than 1) assigned in the transitions matrix in the place of p_{old}, which is its probability value assigned by either the pre-blending or the blending stage. After all new values have been assigned, the transition matrix is normalised so that every row adds to 1. A PIM value of 0 reduces the probabilities of transitions in $\mathtt{A}_{\mathtt{i-j}}$, $\mathtt{B}_{\mathtt{i-X}}$ and $\mathtt{B}_{\mathtt{X-i}}$, resulting in harmonisations that most probably use transitions (and, subsequently, chords) exclusively from one initial harmonic space, making connections between initial spaces less possible. Conversely, a PIM value of 1 encourages harmonisations that transit from $\mathtt{I}_{\mathtt{i}}$ to $\mathtt{I}_{\mathtt{j}}$ and vice versa by using existing chords (transitions in $\mathtt{A}_{\mathtt{i-j}}$) or even new ones (transitions in $\mathtt{B}_{\mathtt{i-X}}$ and $\mathtt{B}_{\mathtt{X-i}}$), producing results that incorporate mixed harmonies of the initial spaces, as well as new chords.

4 Examples of harmonisations using compound harmonic idioms

Concept invention through blending harmonies is a new, currently vague and unexplored field. Therefore, when blending harmonic styles there is no concrete expectation about the result. There are, however, specific problems in harmony that need to be resolved creatively, in which tools like concept invention can be used to propose many alternative and diverse solutions within a unified framework. In this study we focus on customised harmonic settings/examples and illustrate aspects of harmony that can be resolved creatively by using the proposed methodology, while an extended empirical evaluation is presented in Zacharakis et al. (2017). In these examples, we investigate whether the products of the system[7] are within an acceptable range of musical solutions, taking into account the aesthetic contexts of the examples. To the best of our knowledge, there is no methodology for resolving creatively harmonic 'problems' similar to the ones examined in this paper. The Bach chorales are used in many examples since this harmonic style is well understood and gives a clear picture of major and minor tonalities along with a relatively stable tonal centre; it is ideal for examining examples that blend major and minor modes (Section 4.1), key transpositions (Section 4.2) and the introductions of chromaticism (Section 4.3).

Evaluating creative systems is a difficult task since there is no well-established and commonly accepted definition of creativity (e.g. Boden (2004); Wiggins (2006); for a comprehensive discussion see Jordanous (2013), chapter 3). The methods proposed so far either focus on the creative processes *per se* Colton et al. (2011), or on the products of creative processes (Ritchie, 2007; Jordanous, 2013). However, ongoing research (Zacharakis et al., 2017) indicates that the intended purposes of concept blending are successfully accomplished in terms of human perception. For instance, blending major and minor modes (Section 4.1) produces a new idiom that is perceived either in between the original modes, or as novel, intrinsically related, space. Similarly, when blending Bach chorale idiom with standard jazz (Section 4.4), the harmonisations produced by the compound harmonic space are perceived either as in between or as novel space. Furthermore, blending transposed key versions of the same idiom introduces novel creative characteristics to the resulting blended idiom; in the case of Bach chorales (Section 4.3), pilot perceptual results with music university students indicate that the tonal character is altered by the introduction of chromatic characteristics into a predominantly diatonic idiom.

The creative potential of the system is tested on different harmonisation tasks that include harmonisations of melodies using either blends of diverse learned idioms or blends of the same learned idiom with different user-selected parameters. Five short melodies were chosen or created, two original, two taken from a folk music repertoire and one classical music excerpt:

1. Original short melody: This short 8-bar melody deliberately avoids the 3rd and 6th melodic degrees of the C scale, making it 'neutral' regarding its classification as major or minor. It consists of two 4-bar phrases (half cadence - full cadence) that form an 8-bar period.

2. Original short melody: This short 10-bar melody includes tonal shifts between major tonalities one tritone apart (the tritone denotes six steps in the circle of perfect 5ths, which is the largest possible distance between two keys in tonal space). It consists of two phrases: the first 4 bars

[7]For all the examples that follow, if not otherwise explicitly stated, the PIM and RBS values in the presented harmonisations are 0.7 and 0.9 respectively.

define C major tonality and the next 6 bars include a cadence into F♯ major and a return to C major.

3. Scottish folk song: '*Ye Banks and Braes*'. This melody is in the G major pentatonic mode. It is a 16-bar rounded-binary form (aa'ba'), comprising four 4-bar phrases, each concluding with half or full cadence.

4. L. v. Beethoven: The melodic theme (*'Ode to Joy'*) from the final movement of the 9th Symphony, transposed into C major. The melody comprises two 4-bar phrases (half cadence - full cadence) that form an 8-bar period.

5. Greek folk song: *'Apopse ta mesanychta'* (Tonight at midnight). The melody is in D Aeolian mode and comprises two 4-bar phrases, with the first one ending on the 4th melodic degree and the second one on the 1st degree (modal centre).

Seven different musical idioms and some of their blends were used for the harmonisation of the above five melodies, presented in the following list:

- BC major & BC minor: The homophonic tonal harmonic idiom of J. S. Bach chorales.
- JA major & JA minor: Mainstream jazz harmony, as encountered in selected jazz standards from the *Real Book*.
- CN: Greek composer Yannis Constantinidis's 20th-century modal style, as encountered in his '44 Greek miniatures for piano' (Tsougras, 2010).
- HM: Paul Hindemith's 20th-century neo-tonal harmonic idiom, as expressed in his 'Six Chansons' for a capella choir.
- WT: Whole-Tone post-tonal harmony, as encountered in selected excerpts from early 20th-century works.

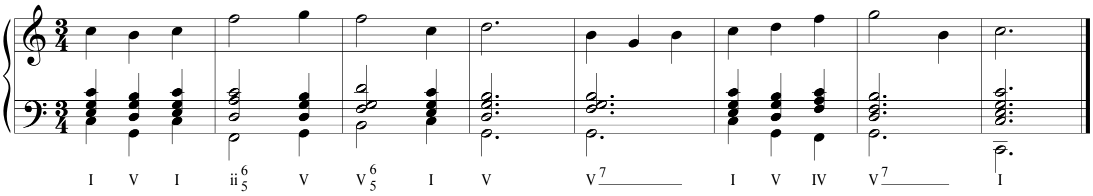
A selection of sixteen harmonisations, taken from a pool of harmonisation attempts with a mixed rate of 'success', are presented in this section[8]. An emphasis was given on the use of tonal idioms (mainly BC major and minor) in the present paper, although the dataset includes numerous other diverse harmonic idioms. This occurs because major-minor tonality is probably one of the most studied harmonic idioms, so it functions as a reference point for testing and demonstrating blending procedures. The system produced raw MIDI files that were further processed by humans using music notation software (Finale 2014). The process involved two stages: 1) correction of music notation issues and enharmonic spellings of pitches in the MIDI file, and 2) manual editing of the produced harmonisation regarding separation of the bass line in a different layer, preservation of a constant number of active voices in the musical texture through the use of octave doublings, and reworking of the inner voices for smoother voice-leading where needed (although a strict application of common-practice voice-leading rules was not pursued). Also, manual analysis of harmonic progressions through the use of Latin roman numeral notation of tonal harmonisation was made in certain cases. The pitch content of the chords was always kept intact, and the bass line was manually altered in very few cases (indicated by * in the scores) in order to avoid stylistic inconsistencies or achieve more effective voice-leading.

4.1 Blending major and minor tonalities

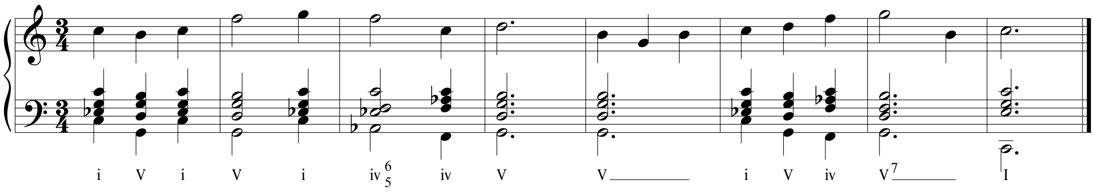
The first example concerns the problem of harmonising a melody with a blended major and minor harmonic style. The short melody is harmonised four times (see Figure 7). It should be noted that apparent parallelisms between the input melody (upper staff) and some voices of the produced harmonisation (lower staff) are not considered errors here, because they merely reflect the property that melody notes (input) are always included in the harmony (output), regardless of voice

[8] Audio files of the presented examples can be found in following address: `https://www.dropbox.com/s/3enwfxe2j9t8ve1/JNMRaudioExamples.zip?dl=0` (temporary address for blind review).

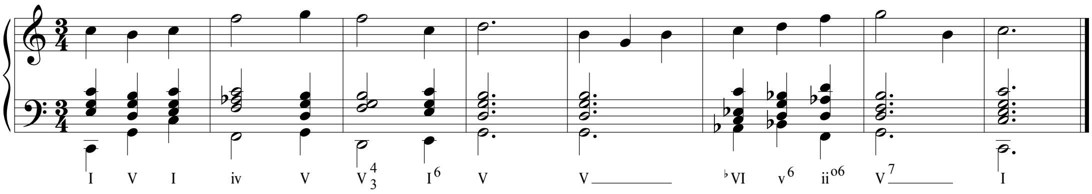
assignment. Further separation of these harmonic 'blocks' into coherent voices following specific voice-leading practices might be handled in a post-processing stage, but is not considered here. The first harmonisation is based on the BC major dataset and uses mainly the I and V harmonic degrees, with sparse use of subdominant-function chords (IV and ii). The second harmonisation uses the BC minor dataset and produces similar results in C minor (use of i, V and iv) with the exception of the final tonic chord, which is major, as encountered in most of the Bach chorales. The first blended harmonisation (low blending parameters) includes a mix of the chords encountered in the two unblended versions (I, V, iv), avoiding the minor tonic (i) and major subdonimant (IV) chords and introducing other chords not used previously (♭VI, v, ii^o). The second blended harmonisation (high blending parameters) uses a different mix of chords from the two parallel tonalities, that includes the minor tonic (i) and the submediant (vi). All four harmonisations conclude with a major tonic, as expected.



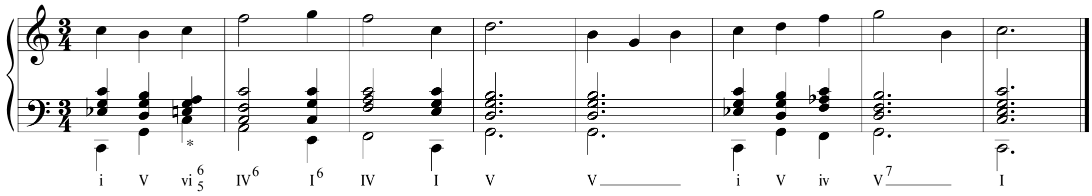
(a) Melody harmonised with the BC major idiom.



(b) Melody harmonised with the BC minor idiom.



(c) Melody harmonised with the blended major-minor BC space, using low PIM and RBS values.

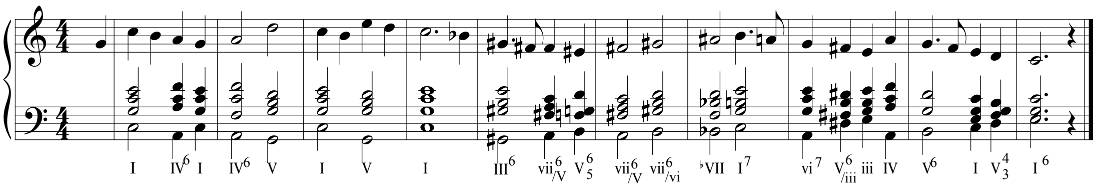


(d) Melody harmonised with the blended major-minor BC space, using high PIM and RBS values.

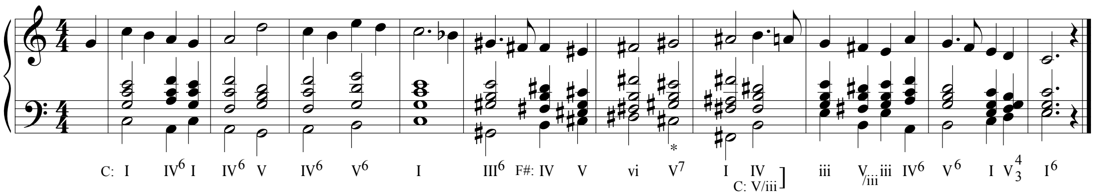
Figure 7: A melody harmonised with idioms learned from a dataset of (a) major (BC major), (b) minor (BC minor) Bach chorales and their harmonic blend with (c) low PIM and RBS values (0.05 and 0.1 respectively) and (d) high PIM and RBS values (0.7 and 0.9 respectively)

4.2 Blending different major keys for modulation purposes

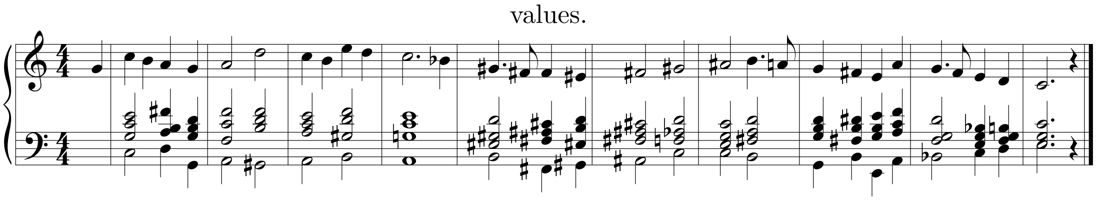
Compound harmonic spaces created through transition blending can also be used for key transpositions, by combining the spaces of the incorporated keys. The short melody is harmonised three times (see Figure 8), all with the use of the BC major (Bach Chorales major) harmonic idiom. The first harmonisation does not incorporate any blending of keys, so the harmoniser attempts to assign chords without modulating away from C major. The result reveals that the melodic shift towards F♯ major has been ignored, however the system has managed to assign chords to the melody's chromatic pitches, albeit with functionally awkward or ambiguous results (G♯ has been harmonised with E major chord, F♯ with F♯ diminished chord, G♯ with G♯ diminished chord and A♯ enharmonically with B♭ major chord, see Figure 8 (a) for an analytical attempt with Roman numerals). The second harmonisation uses the blended C-F♯ major space with low PIM and RBS values, so the system is now able to identify the modulating segment of the melody, and manages to suggest functionally correct chords for both the shift towards F♯ and the return to C major, as the harmonic analysis reveals (see Figure 8 (b)). All the chords are triads, except for the dominant, which appears with its 7th in two cases. The third harmonisation uses the blended space again, but with high PIN and RBS values. The result is quite original, as apart from the modulation, which has been identified and harmonised correctly, the system introduces chromaticism within each tonal region, with unexpected assigned chords in several cases (e.g. the B minor 7th in b. 1, the D half-diminished in b. 6, the G minor 7th and C major 7th in b. 9). This harmonisation displays elements of unexpected originality to an extend that an explicit functional harmonic analysis would be unsuitable, so it has been avoided.



(a) Melody harmonised with the C major BC space.



(b) Melody harmonised with the blended C and F♯ BC major spaces, using low PIM and RBS values.



(c) Melody harmonised with the blended C and F♯ BC major spaces, using high PIM and RBS values.

Figure 8: A melody harmonised with the learned BC major idiom with (a) C major tonality and blended versions of BC major in the tonality of C and its transposition in F♯ major with (b) low PIM and RBS values (0.05 and 0.1 respectively) and (c) high PIM and RBS values (0.7 and 0.9 respectively).

4.3 Creative blending of different major keys

Combining harmonic spaces in different keys can also be used creatively, e.g. for introducing chromatic harmonic characteristics when harmonising even simple tonal melodies, without necessarily focussing on the problem of key transposition. *'Ye Banks and Braes'* is harmonised four times (see Figure 9), all with the use of the BC major harmonic idiom. The first harmonisation does not implement blending and the space used is the G major tonality. The obtained result is a functionally correct and original tonal harmonisation, featuring half and perfect cadences in b. 4, 8, 16, main use of the three basic harmonic degrees (I, IV, V) and chromatic elements of tonicisations (V7/IV, V7/ii) (see analysis in Figure 9 (a)). The other three harmonisations use the blended spaces G major-B♭ major (3 semitones), G major-B major (4 semitones) and G major-C♯ major (6 semitones) respectively and feature forced chromaticism applied on the pentatonic/diatonic space defined/implied by the melody. Harmonic analysis has not been included for these three examples, although it was obtainable in most cases, because of instances where the harmony deviated from functional progressions towards free chromaticism. Some interesting cases of chords are worth presenting though. For the G major-B♭ major space: the G half-diminished chord in b. 5 and the linear chromatic progression in b. 13. For the G major-B major space: the augmented 6th chord in b. 5 and the chromatic linear progressions in b. 6-7 and b. 11. For the G major-C♯ major space: the unexpected chromatic beginning in F♯ minor in b. 1-2 and the linear chromatic progression in b. 6 and b. 14. Since the melody is purely pentatonic and does not even imply chromaticism, it is interesting that the blended tonal spaces can produce such a diverse range of forced harmonic chromaticism, with elements of tonal mixture, chords of ambiguous functionality, and chromatic contrapuntal chords. What is equally interesting is that the produced chords cannot always be explicitly identified as belonging to one of the blended spaces.

4.4 Creative blending of Bach chorale with jazz standard styles

Blending can be used creatively for combining two idioms from different eras. The '*Ode to joy*' melody is harmonised three times (see Figure 10). The first harmonisation uses the BC major idiom (no blending) and consists of the alternation of only two triads: the tonic and the dominant in root position and without 7th extensions (Figure 10 (a)). The second uses the JA idiom (no blending also) and conforms to the mainstream jazz harmony rules: all chords include major or minor 7ths, the main harmonic pattern is the ii-V-I turnaround and there is a tonicisation of the IV at the beginning of the second phrase instead of a half cadence (Figure 10 (b)). The third harmonisation is based on the harmonic idiom produced by the blending of the previous two. As shown in Figure 10 (c), there is a mix of simple and extended triads and a combination of tonicisations and progressions in the circle of 5ths. Interestingly, now the tonicisation of IV occurs in b. 2-4 through a turnaround, a tonicisation of ii is prepared but avoided in b. 5, and a chromatic tonicisation of vi is observed in b. 7. These elements were not part of the unblended versions and seem implicitly only related to either idiom, although there is a sense that the jazz idiom dominates the system's choices.

4.5 Creative blending of diverse harmonic spaces

Even more 'creative' harmonisations are produced by the system when harmonies of diverse and idiosyncratic idioms are blended. The Greek folk melody is harmonised two times (see Figure 11), using blends of diverse, mainly post-tonal, harmonic idioms. The first harmonisation uses a blend of Yannis Constantinidis's harmonic style (20th-century chromatic modal harmony, see Tsougras (2010)) and Whole-Tone harmony. It seems that Constantinidis's harmony dominates, with its parallel voice-leading (b. 1, 3), free use of minor or major 7th chords, and conclusion on an open-5th sonority (b. 8), however a characteristic influence of the WT space can be observed in b. 5 (WT chord C-D-E-G♯). The second harmonisation is based on a blending between minor jazz harmony (extended tonal idiom) and the neo-tonal harmonic idiom of Hindemith (free chromatic harmony with pitch centres). The chords suggested by the system are mainly extended triads with loose harmonic functions, but there are some notable exceptions, either as free mildly dissonant chords (mostly free verticalisations of diatonic sets), e.g. the A-D-G-B sonority in b. 2 and 4, and the quartal

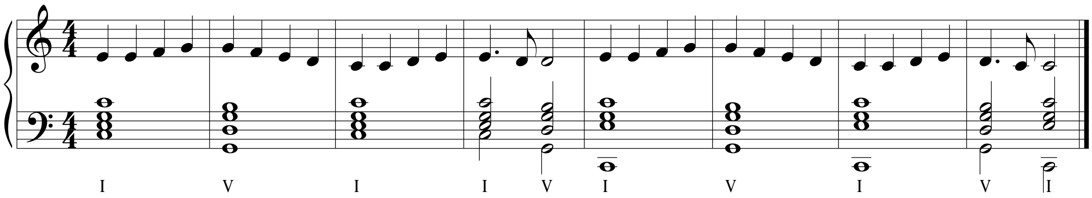
(a) Traditional Scottish melody harmonised with the G major key of the BC space.

(b) Traditional Scottish melody harmonised with the blended keys of G and B♭ major.

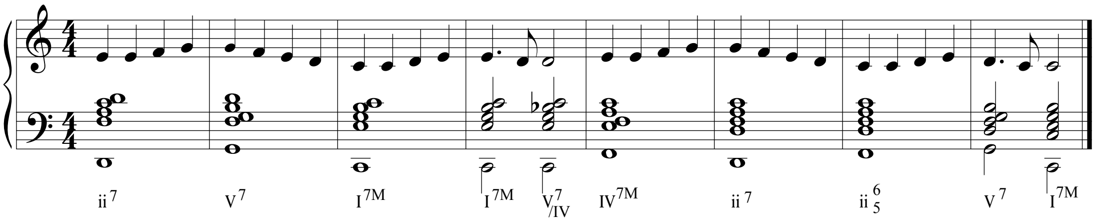
(c) Traditional Scottish melody harmonised with the blended keys of G and B major.

(d) Traditional Scottish melody harmonised with the blended keys of G and C♯ major.

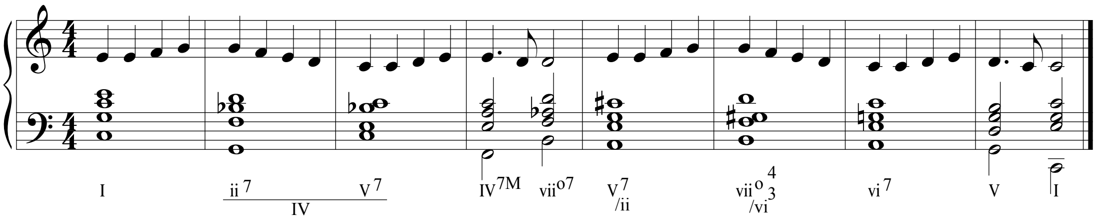
Figure 9: A traditional Scottish melody harmonised with the learned BC major space in (a) the G major key and with blended versions of BC major in the keys of (c) G major and B♭ major, (b) G major and B major and (c) G major and C♯ major.



(a) The melody of *Ode to joy* harmonised with the BC major space.



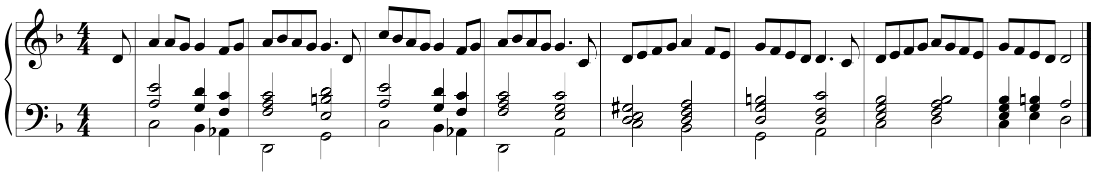
(b) The melody of *Ode to joy* harmonised with the JA major space.



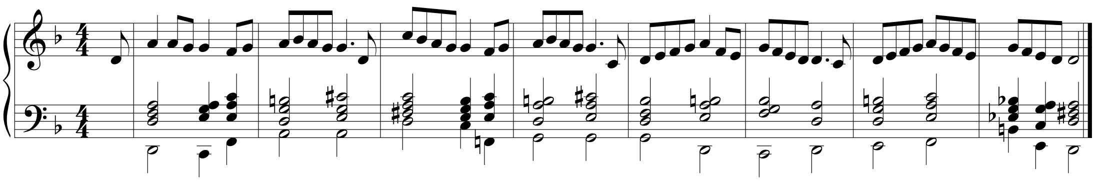
(c) The melody of *Ode to joy* harmonised with the the blended Bach chorale and jazz standards styles.

Figure 10: Beethoven's *Ode to Joy* theme harmonised in the style of (a) BC, (b) jazz standards and (c) their combined harmonic space.

chords D-E-A-B (b. 5) and C-F-G-B♭ (b. 6) or as the highly dissonant sonority B-E♭-G-B♭ in b. 8. However, although certain elements of the harmonisations may be classified as stemming from one of the blended idioms, the overall produced idioms feel original and coherent.



(a) A traditional Greek melody harmonised in the harmonic style of Constantinides blended with the whole-tone harmony.



(b) A traditional Greek melody harmonised in the harmonic style of jazz standards blended with the Hindemith's harmonic idiom.

Figure 11: A traditional Greek melody harmonised with the blended harmonic styles of (a) Constandinidis with whole-tone and (b) jazz standards with Hindemith.

5 Conclusions

Creative melodic harmonisation through the combination of harmonic spaces is examined in this paper; conceptual blending is integrated into a Markov chord transition methodology that is part of an idiom-independent melodic harmonisation assistant that learns from harmonic data. The algorithmic framework for conceptual blending developed in the context of the COINVENT (Concept Invention Theory) project is utilised to blend transitions of chord (pairs of successive chords), referred to as the 'input transitions', between different harmonic idioms, producing new blended transitions. Transition blending is based on combining a priori defined transition features, while the best blends are identified through a process that ranks all resulting blends according to the salience of their incorporated features; the salience of features is automatically induced by statistical assessment on the learned input idioms. The best blends of the most usual transitions in two initial harmonic idioms are then used to construct a new 'compound' harmonic space, that includes the chords and the allowed transitions of the initial idioms, along with new chords and new transitions.

The creative harmonic capabilities of the system have been examined under many blending melodic harmonisation settings, revealing a number of different previously unexplored cases (in the context of musical artificial intelligence) where this methodology can be applicable – from robust problem solving to purely experimental harmonic exploration. The examples presented in this paper discuss some interesting application of the proposed methodology for blending: (i) major and minor modes; (ii) different major keys for modulations; (iii) different major keys for increasing chromaticism; (iv) different harmonic styles, e.g. Bach chorales and jazz standards; and (v) diverse harmonic spaces for exploring novel harmonic ideas. Therefore, the presented system could on the one hand be used for providing conventional harmonic solutions, constituting a useful tool, e.g., for music education. On the other hand, it can be used for generating unconventional harmonic material, providing composers with a tool that can produce many creative alternatives in harmonising a melody. Additionally, this system could be used by non-experts in music for experimenting in different blending combinations, or it could be also integrated in other software, e.g. games, for producing novel, unique and diverse musical backgrounds. The new possibilities that the proposed system offers highlight the overall new capabilities that are introduced in computational creativity by conceptual blending, as algorithmically approached in the context of the COINVENT project.

Thorough evaluation of how the products of this system are perceived is an ongoing research, but pilot results (Zacharakis et al., 2017) indicate that the intended purposes of blending are met, with the system creating compound idioms that are perceived either as in between the blended ones, or as something completely new, yet related to the original ones. An interesting future research direction is towards increasing the self-awareness of the system, by developing methods that automatically categorise the products of the system, either by performing style classification or qualitative evaluation. It would be also interesting to examine how current state-of-the-art algorithms for style classification would classify blended harmonisations (as belonging to one of the input spaces, in-between them, or as belonging to a whole new style). Self-awareness on this level would allow the system to re-adjust the blending parameters, i.e. PIM and RBS, so that more meaningful blended harmonisations are produced without user intervention.

Acknowledgements

This work is supported by the COINVENT project (FET-Open grant number: 611553).

References

Antovic, M. (2011). Musical metaphor revisited: Primitives, universals and conceptual blending. *Universals and Conceptual Blending (February 17, 2011)*.

Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. Psychology Press.

Boden, M. A. (2009). Computer Models of Creativity. *AI Magazine*, 30(3):23.

Cambouropoulos, E. (2012). A Directional Interval Class Representation of Chord Transitions. In *In Proceedings of the Joint Conference ICMPC-ESCOM 2012 (12th International Conference for Music Perception and Cognition, & 8th Conference of the European Society for the Cognitive Sciences of Music)*, ICMPC-ESCOM 2012.

Cambouropoulos, E., Kaliakatsos-Papakostas, M., and Tsougras, C. (2014). 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.

Cambouropoulos, E., Kaliakatsos-Papakostas, M., and Tsougras, C. (2015). 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)*.

Cambouropoulos, E., Katsiavalos, A., and Tsougras, C. (2013). Idiom-independent harmonic pattern recognition based on a novel chord transition representation. In *In Proceedings of the 3rd International Workshop on Folk Music Analysis (FMA2013)*, FMA 2013.

Chordia, P., Sastry, A., Mallikarjuna, T., and Albin, A. (2010). Multiple viewpoints modeling of tabla sequences. In *ISMIR*, volume 2010.

Cleary, J. G. and Teahan, W. (1995). Experiments on the zero frequency problem. In *Proc. Data Compression Conference*, volume 480.

Colton, S., Pease, A., and Charnley, J. (2011). Computational Creativity Theory: The FACE and IDEA Descriptive Models. In *2nd International Conference on Computational Creativity*.

Confalonier, R., Schorlemmer, M., Plaza, E., Eppe, M., Kutz, O., and Peñaloza, R. (2015). 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*.

Cook, N. (2001). Theorizing musical meaning. *Music Theory Spectrum*, 23(2):170–195.

Eppe, M., Confalonier, R., Maclean, E., Kaliakatsos-Papakostas, M., Cambouropoulos, E., Schorlemmer, M., Codescu, M., and Kühnberger, K.-U. (2015a). Computational invention of cadences and chord progressions by conceptual chord-blending. In *International Joint Conference on Artificial Intelligence (IJCAI) 2015*.

Eppe, M., Maclean, E., Confalonieri, R., Kutz, O., Schorlemmer, M., and Plaza, E. (2015b). ASP, Amalgamation, and the Conceptual Blending Workflow. In Calimeri, F., Ianni, G., and Truszczynski, M., editors, *Logic Programming and Nonmonotonic Reasoning*, volume 9345 of *Lecture Notes in Computer Science*, pages 309–316. Springer International Publishing.

Fauconnier, G. and Turner, M. (2003). *The Way We Think: Conceptual Blending And The Mind's Hidden Complexities*. Basic Books, New York, reprint edition edition.

Goguen, J. (2006). Mathematical Models of Cognitive Space and Time. In Andler, D., Ogawa, Y., Okada, M., and Watanabe, S., editors, *Reasoning and Cognition*, volume 2 of *Interdisciplinary Conference Series on Reasoning Studies*. Keio University Press.

Goguen, J. and Harrell, D. F. (2010). Style: A Computational and Conceptual Blending-Based Approach. In Argamon, S. and Dubnov, S., editors, *The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning*, pages 147–170. Springer, Berlin.

Granroth-Wilding, M. and Steedman, M. (2014). A robust parser-interpreter for jazz chord sequences. *Journal of New Music Research*, 0(0):1–20.

Jordanous, A. K. (2013). *Evaluating computational creativity: a standardised procedure for evaluating creative systems and its application*. PhD thesis, University of Sussex.

Kaliakatsos-Papakostas, M. and Cambouropoulos, E. (2014). 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.

Kaliakatsos-Papakostas, M., Cambouropoulos, E., Kühnberger, K.-U., Kutz, O., and Smaill, A. (2014). Concept Invention and Music: Creating Novel Harmonies via Conceptual Blending. In *In Proceedings of the 9th Conference on Interdisciplinary Musicology (CIM2014)*, CIM2014.

Kaliakatsos-Papakostas, M., Confalonier, R., Corneli, J., Zacharakis, A., and Cambouropoulos, E. (2016a). An argument-based creative assistant for harmonic blending. In *Proceedings of the 7th International Conference on Computational Creativity (ICCC)*, page (accepted).

Kaliakatsos-Papakostas, M., Makris, D., Tsougras, C., and Cambouropoulos, E. (2016b). Learning and creating novel harmonies in diverse musical idioms: An adaptive modular melodic harmonisation system. *Journal of Creative Music Systems*, 1(1):(In press).

Kaliakatsos-Papakostas, M., Zacharakis, A., Tsougras, C., and Cambouropoulos, E. (2015). 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.

Koops, H. V., Magalhães, J. P., and de Haas, W. B. (2013). A functional approach to automatic melody harmonisation. In *Proceedings of the First ACM SIGPLAN Workshop on Functional Art, Music, Modeling & Design*, FARM '13, pages 47–58, New York, NY, USA. ACM.

Makris, D., Kaliakatsos-Papakostas, M., and Cambouropoulos, E. (2015). A probabilistic approach to determining bass voice leading in melodic harmonisation. In *Mathematics and Computation in Music: Proceedings of the 5th International Conference, MCM 2015*, page Accepted for publication, London, UK. Springer, Berlin.

Moore, A. F. (2013). *Song means: Analysing and interpreting recorded popular song.* Ashgate Publishing, Ltd.

Ontañón, S. and Plaza, E. (2012). On knowledge transfer in case-based inference. In Agudo, Belén, D. and Watson, I., editors, *Case-Based Reasoning Research and Development*, volume 7466 of *Lecture Notes in Computer Science*, pages 312–326. Springer Berlin Heidelberg.

Ox, J. (2014). Analogy and conceptual blending are part of a visualization toolkit for artists and scientists: Introducing the cognitive space transfer. *Methods*, 1:6.

Pachet, F., Roy, P., and Barbieri, G. (2011). Finite-length markov processes with constraints. *transition*, 6(1/3).

Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17(1):67–99.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1):35–53.

Schorlemmer, M., Smaill, A., Kühnberger, K.-U., Kutz, O., Colton, S., Cambouropoulos, E., and Pease, A. (2014). Coinvent: Towards a computational concept invention theory. In *5th International Conference on Computational Creativity (ICCC) 2014.*

Spitzer, M. (2004). *Metaphor and musical thought.* University of Chicago Press.

Stefanou, D. and Cambouropoulos, E. (2015). Enriching the blend: Creative extensions to conceptual blending in music. In *Proceedings of the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM).*

Tsougras, C. (2010). The Application of GTTM on 20th Century Modal Music: Research Based on the Analysis of Yannis Constantinidis's "44 Greek Miniatures for Piano". *Musicae Scientiae*, 14(1 suppl):157–194.

Tsougras, C. and Stefanou, D. (2015). Conceptual Blending and Meaning Construction: A Structural/hermeneutical Analysis of the 'Old Castle' from Musorgsky's 'Pictures at an Exhibition'. In *Proceedings of the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM).*

Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7):449–458.

Zacharakis, A., Kaliakatsos-Papakostas, M., and Cambouropoulos, E. (2015). Conceptual blending in music cadences: A formal model and subjective evaluation. In *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2015)*, Malaga, Spain.

Zacharakis, A., Kaliakatsos-Papakostas, M., Tsougras, C., and Campouropoulos, E. (2017). Empirical methods for evaluating music structural blends: a case study on melodic harmonisation. *Musicae Scientiae*, page (forthcoming).

Zbikowski, L. M. (2002). *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. Oxford University Press.

Zbikowski, L. M. (2008). Metaphor and music. *The Cambridge handbook of metaphor and thought*, pages 502–524.

SUBMITTED DRAFT - UNDER REVISION

Creating musical cadences via conceptual blending: Empirical evaluation and enhancement of a formal model

Asterios Zacharakis, Maximos Kaliakatsos-Papakostas, Costas Tsougras & Emilios Cambouropoulos

Aristotle University of Thessaloniki, Thessaloniki, Greece.

Abstract

The cognitive theory of conceptual blending may be employed to understand the way music becomes meaningful, but, at the same time it may form a basis for musical creativity *per se*. This work constitutes a case study whereby conceptual blending is used as a creative tool for inventing musical cadences. Specifically, the Perfect Authentic and the renaissance Phrygian cadences are used as input spaces to a cadence blending system that produces various cadential blends based on musicological and blending optimality criteria. A selection of 'novel' cadences is subject to empirical evaluation in order to gain a better understanding of perceptual relationships between cadences. Pairwise dissimilarity ratings between cadences are transformed into a perceptual space and a verbal attribute magnitude estimation method on six descriptive axes (preference, originality, tension, closure, expectancy and fit) is used to associate the dimensions of this space with descriptive qualities (closure and tension emerged as most prominent qualities). The novel cadences generated by the computational blending system are mainly perceived as one-sided blends (i.e. blends where one input space is dominant), since categorical perception seems to play a significant role (especially in relation to the upward leading note movement). Insights into perceptual aspects of conceptual bending are presented and ramifications for developing sophisticated creative systems are discussed.

Keywords: conceptual blending, musical cadence, computational creativity, empirical evaluation, harmony perception

Introduction

New concepts may be created either by 'exploring' previously unexplored regions of a given conceptual space (exploratory creativity) or by transforming established concepts in novel ways (transformational creativity) or by making associations between different conceptual spaces that share some structural relations (combinational creativity) – Boden maintains that the latter, i.e., combinational creativity, has proved to be the hardest to describe formally (Boden, 2009). This paper explores aspects of combinational creativity in the domain of music, and more specifically, the harmonic structure of music.

Conceptual blending is a cognitive theory developed by Fauconnier and Turner (2003) whereby elements from diverse, but structurally-related, mental spaces are 'blended', giving rise to new conceptual spaces that often possess new powerful interpretative properties allowing better understanding of known concepts or the emergence of novel concepts altogether. Conceptual blending theory is useful for explaining the cognitive process that humans undergo when engaged with creative acts, and is akin to Boden's notion of combinational creativity. A computational framework that extends Goguen's formal approach (Goguen 2006) has been developed in the context of the COINVENT (Concept Invention Theory) project (Schorlemmer et al., 2014). According to this framework, two *input spaces* are described as sets of weighted properties and relations, and after their *generic space* is computed, the *amalgamation* process (Eppe et al., 2015) leads to the creation of consistent blends that are optimal according to some criteria relating to the blending process and to the knowledge domain of the modeled spaces (the amalgamation process potentially includes multiple 'generalization paths', leading to many different blends).

With regard to music, conceptual blending has been predominantly theorized as the cross-domain integration of music structural and extra-musical domains such as text or image (e.g. Zbikowski, 2002 & 2008; Cook, 2001; Moore, 2012). Additionally, it has been studied in the context of 'musicogenic' meaning (Koelsch, 2013), which refers to physical, embodied, emotional and personality-related responses to music; such studies include work on music and motion by Johnson and Larson (2003) or empirical studies on pitch perception and image schemata in children by Antovic (2009, 2011). Finally, there have been studies that touch upon issues of structural mappings/blending between different spaces within the music structure domain *per se* (such as mappings between incongruous tonalities (Ox, 2014) and different tonal pitch space theories (Spitzer, 2003)). Almost all of the above studies examine conceptual blending in retrospect, analyzing and explaining existing metaphors/blends rather than taking a bottom-up, creative perspective of *generating* novel blends. A more extended discussion and critical examination of conceptual blending processes in music is presented by Stefanou and Cambouropoulos (2015).

In this paper it is maintained that the creative potential of conceptual blending (i.e., invention of new blends) in the domain of music is, probably, most powerfully manifested in processes that enable structural blending. To substantiate this potential, a proof-of-concept autonomous computational creative system that performs melodic harmonization is being developed (Kaliakatsos-Papakostas et al., 2016). A core component of this system is a transition blending mechanism that has been applied, among other things, to well defined harmonic concepts such as harmonic cadences (Eppe et al., 2015; Zacharakis et al., 2015b). The present work focuses on conceptual blending of musical cadences (with well-established functional/voice-leading characteristics) and reports in detail algorithmic and empirical findings that relate to its application. The particular focus on cadences comes from the fact that they constitute a most salient harmonic concept and are of major importance in tonal music. The cadence's significance lies not only in its form-creating function, i.e. the delineation of phrase/group boundaries that give rise to hierarchical grouping structure, but, also, in that its harmonic content contributes, to a considerable extent, to the special character of a harmonic idiom in which it functions as an indispensable closure element (e.g. Bigand & Parncutt, 1999; Huron, 2006, Chapter 9; Sears, 2015; Caplin, 2004; Aldwell & Schachter, 1999). The insight obtained by this proof-of-concept approach will be exploited to develop a system capable of performing harmonic blending between different musical idioms in a melodic harmonization task.

The blending methodology is applied to two distinct musical cadences: the tonal Perfect Authentic Cadence (PAC), as encountered in 18th and 19th century tonal music and the modal Phrygian cadence, as encountered in 16th century (Renaissance) modal music (Figure 1). The perfect authentic cadence is described as a functional dominant-to-tonic chord progression (Sears, 2015; Aldwell & Schachter, 2003; Caplin, 1998) constituted from a $V^{(7)}$ chord in root position –prepared by a chord with pre-dominant function– leading to a I chord in root position and with the tonic in the upper voice (^1). The three- or four-voice Phrygian cadence is described as a contrapuntal progression (Barnett, 2002; Schubert, 1999; Collins Judd, 2002) based on a two-voice linear movement and constituted from a $\flat$ vii^6 chord leading to a I or i or I_{omit3} chord with the tonic in the upper voice (^1) (see Figure 1).

For the purposes of blending, the cadences are modeled as rich concepts that embody several properties. Thus, the above two cadences are represented not only as chord types but, additionally, as collections of notes and note transitions with weights attached to each note or note transition based on functional/voice-leading properties, such as semitonal resolution of the leading note, type of harmonic progression expressed as distance between chordal roots, the existence of the tritone in the penultimate chord of the perfect cadence (lines in Figure 1 indicate important notes and note transitions). For instance, in the perfect cadence, the upward leading note is probably the most salient component of the penultimate chord, as it appears in all the main dominant chord types (V, V^7 and vii^o), the root and the seventh of the dominant chord are salient but hypothetically less so than the leading note, and the fifth of the dominant is the least important component, as it can be omitted in most cases. Accordingly, in the Phrygian cadence, the most characteristic component of the penultimate chord is the downward leading note (downward semitone movement: ^2 to ^1), the root is salient but hypothetically less so (upward whole tone movement: ^7 to ^8) and the fifth is the least salient element of the chord. In short, the most prominent characteristics of the two cadences are assumed to be the upward leading note of the perfect and the downward leading note of the Phrygian cadence. The two input spaces (perfect and Phrygian) are represented as being equally important in the blending process; however, we expect the perfect cadence to be more prominent as a cadential schema in the mind of contemporary listeners, due to their comparatively longer exposure to classical tonal music rather than to Renaissance modal music (this is examined in the perceptual experiments below).

(Figure 1)

Applying the proposed conceptual blending system (see next section) to the perfect and Phrygian input spaces, the tritone substitution progression (see Figure 1) emerges; this cadence is highly ranked by the proposed blending process as it incorporates most of the salient features of both cadences (it includes both the most salient upward and downward leading notes). It is worth noting that the computational system 'invents' this cadential type, which emerged in jazz, centuries after the main tonal/modal input cadences. On the other hand, a cadence that does not include any of these two properties, such as the backdoor progression (also used in jazz), may also appear as a blend (depending on how blends are rated/selected), but much lower in the ranking. Many other blends are possible, seven of which are further examined empirically.

Given that our computational system is capable of inventing novel cadential schemata by blending basic cadences, we are particularly interested in the following questions: Are the novel cadences generated by the system perceived as being one-sided blends (i.e., closer to one of the input cadences) or are they balanced double-sided blends (in between the perfect and Phrygian cadences)? Are the generated highly-ranked new cadences perceived by listeners as being 'good' blends between the perfect and Phrygian cadences (in case of double-sided blends) or as being interesting new versions of the perfect or Phrygian cadences (in case of single-sided blends)? How do listeners perceive the new cadences in terms of originality, expectancy, sense of closure and tension? Which cadences do they prefer? The current study attempts to address these issues through a series of subjective experiments. It gives no definitive answers, but hopefully the descriptions, experiments and discussions below will shed some light into perceptual aspects of musical creativity, opening the way for more extensive and thorough studies in the future.

Evaluating creativity - either human or computational - is a non-trivial task, especially when the assessment of aesthetic quality is also involved. The matter is further complicated by the fact that the mere definition of creativity is problematic and not commonly accepted as many authors approach it from different perspectives (e.g. Boden, 2004; Wiggins, 2006; for a comprehensive discussion see Jordanous, 2013, chapter 3). As a result, creativity is often broken down into partial constituent dimensions (e.g. novelty, value, surprise, problem solving ability, originality, divergence, etc.) (e.g. Maher et al., 2013; Jordanous, 2013). In terms of assessing a creative system, the two usual approaches are either to directly evaluate the product of the system or to evaluate the production mechanism (Pearce and Wiggins, 2001). The former can be also viewed as a summative evaluation (Jordanous, 2013, chapter 1) whereby the overall creativity of a system is sought for. The latter is a formative evaluation process whose objective is to provide evaluation feedback concerning certain attributes of the creative system during the development stage and, thus, direct possible improvements. The present work adopts essentially the summative approach (evaluation of the end products of the system) but, also, takes into account the formative characteristics of the creative system with a view to increasing its creative potential.

The empirical evaluation was performed by means of two subjective tests: a main non-verbal dissimilarity rating listening test (following a preliminary study reported in Zacharakis et al. 2015b) and a complementary verbal subjective test. In the main experiment we opted for a (non-verbal) pairwise dissimilarity rating listening test between nine cadences (the two originals and seven blends). We subsequently applied Multidimensional Scaling (MDS) analysis to the acquired data and used the produced spatial configuration as an indirect way to measure the relation of blends to the input cadences. One intuitive assumption is that an 'ideal' double-sided blend should be 'equally' similar to each of the input cadences (it should resemble both input concepts) and therefore should appear in between them (ideally near the middle), while weaker one-sided blends should be positioned closer to either of the originals (off the middle).

In a complementary experiment, a descriptive type of subjective evaluation (Verbal Attribute Magnitude Estimation) was employed to assess qualities of the produced blends. In this experiment, the nine cadences were presented to listeners in two different harmonic contexts, namely, a tonal minor context and a Phrygian context, resulting in 18 cadential stimuli. Listeners were asked to rate each cadence according to preference, degree of tension, closure effect, originality, expectedness and fit within the corresponding tonal/modal context. It was hypothesized that within a certain context some cadences might be considered more original/unexpected than other blends, evaluating thus directly certain aspects of the creative blending system.

In the first section below, a systematic description of the conceptual blending mechanism is presented along with a formal representation of cadences. The next two sections present and discuss the two empirical experiments. An overall discussion of the findings concludes the paper.

A computational method for conceptual blending: inventing new cadences

The intended goal of a computational system for conceptual blending is to achieve a combination of different structural parts of two input conceptual spaces so that the generated blended space encompasses new structure and novel properties, preserving at the same time the common parts of the inputs. In computational creativity, conceptual blending has been modeled by Goguen (2006) as a generative mechanism, according to which input spaces are modeled as *algebraic specifications* and a blend is computed as a categorical *colimit*. A computational framework that extends Goguen's approach has been developed in the context of the COncept INVENtion Theory[1] (COINVENT) project (Schorlemmer et. al., 2014) based on the notion of *amalgams* (Ontañón and Plaza, 2010; 2012). According to this framework, *input spaces* are described as sets of *properties,* and an *amalgam-based* workflow (Confalonieri et al., 2015; Eppe et al., 2015) finds the blends by generalizing (or removing) input properties until a *generic space* (i.e., the set of common properties between the input spaces) is found; intermediate generalized versions of the input spaces are ‘merged’ to create blends that are consistent or satisfy certain properties related to the knowledge domain (see Figure 2)[2].

In this paper the specific case of blending the perfect and Phrygian cadences discussed above is examined. For simplicity, we assume that each cadence consists of two chords, the second of which is always a Cm; only the penultimate chord can be altered through blending. The properties that are used for describing a cadence concern either its penultimate chord or pitch class differences/intervals between the two constituent chords (described later in Table 1). When blending two cadences, the amalgam-based algorithm first computes their generic space (common properties illustrated as point 1 in Figure 2). After the generic space is found for two given input cadences, the amalgam-based process attempts to compute their *amalgam*, which is the *unification* of their content. If the resulting amalgam is inconsistent, then it iteratively generalizes the properties of the inputs (point 2 in Figure 2), until the resulting unification is consistent (point 3 in Figure 2). For instance, trying to directly unify the transitions I_1: G7 → Cm and I_2: B♭m → Cm would yield an inconsistent amalgam, since a transition cannot both include and *not* include a leading note to the tonic (which are properties of I_1 and I_2 respectively). Therefore, the amalgam-based process generalizes the property that creates the clash in one of the inputs (e.g., the property describing the absence of leading note would be left empty in I_2) and tries to unify the generalized versions of the inputs again.

After a number of generalization steps are applied (point 2 in Figure 2), the input spaces are generalized ‘enough’ so that the resulting blend is consistent (point 3 in Figure 2). However, it may be the case that the blend is *not complete*, in the sense that this process may have generated an over-generalized result by over-generalizing the inputs during the amalgamation step. Blends are then completed by *blending completion* (Fauconnier & Turner, 2003), which is a domain-specific process that uses background knowledge to consistently assign specific properties to generalized terms. For instance, in the hitherto examined case, blend completion is used for completing the A♭ note (which does not exist in any input) as the fifth of the penultimate chord when obtaining the tritone substitution cadence (Figure 1).

(Figure 2)

A formal description of cadences for generative conceptual blending

A cadence is described by several properties that concern both the penultimate chord and musical values that change during its transition to the final chord - these properties are shown in Table 1. Among the properties that are included in the description of the penultimate chord are its root and type; chord roots are necessary for computing the root difference with the final chord. For computing the root and type in a consistent manner for all utilized chords, the *General Chord Type (GCT)* representation (Cambouropoulos et al., 2014) has been employed, which allows the re-arrangement of the notes of a harmonic simultaneity such that abstract types of chords along with their root may be derived. This encoding is inspired by the standard roman numeral chord type labeling, but is more general and flexible since it can be used to describe chords in any musical idiom. The GCT algorithm finds the maximal subset of notes in a simultaneity that contains only consonant intervals, given a user-defined consonance-dissonance classification of intervals that reflects sensory and/or culturally-dependent notions of consonance/dissonance. This maximal subset forms the *base* upon which the chord type is built, while the lowest note of the base is the *root* of the chord; any remaining notes that cannot be a part of the maximally consonant subset are included in the *extension* of the GCT type. For example, by considering the unison, third/sixth and perfect fourth/fifth intervals as consonant, the GCT representation of the first degree (I) chord in a major scale is [0, [0 4 7]], where 0 indicates the root note in relation to the scale (0 is the scale's first degree) and [0 4 7] is the chord's type (4 indicates a major third and 7 a perfect fifth). Accordingly, a V7 chord is denoted by [7, [0 4 7], [10]], where 10 is the extension (minor seventh), which cannot be included in the base considering that the tritone and minor seventh intervals are dissonant. As the GCT representation is general and can be applied to non-standard tonal systems such as modal harmony and, even, atonal harmony, the blending scheme considered for the cadences described herein, can be generalized to cadences of practically any musical idiom.

(Table 1)

Cadence properties 1-3 describe the first (penultimate) chord of the cadence; the first two properties (chord root and type) are computed by the GCT algorithm. The pitch classes of the chord are described in Property 3. Property 4, which is the difference between the chord roots, is an integer between -5 and 6 that indicates the pitch class difference between the roots of the first and second chords of the cadence. Property 5 captures the existence of a common pitch class between the two chords, while properties 6 and 7 indicate the existence of a semitone movement (upward and downward respectively) in any pitch class of the cadence transition. Properties 5, 6 and 7 actually indicate if there is a 0, 1 or -1 in the Directional Interval Class (DIC) (Cambouropoulos, 2012; Cambouropoulos et al., 2013), flagging whether there are small pitch class voice leading movements (repeating notes or semitone movements) in the cadence. Properties 8 to 10 are used to indicate whether there is a semitone movement (property 10) to the tonic from the first to the second chord of the cadence, as well as whether this movement is ascending (property 8) or descending (property 9); these properties reflect the importance of the leading note (upwards or, even, downwards).

Generating, rejecting and ranking blends

Table 2 illustrates a blending example, where the tritone substitution cadence is created from the perfect and the Phrygian cadences. This blend incorporates properties from both input spaces, many of which are common to both spaces, while new properties have also been added through completion. Specifically, this blend includes four properties exclusively from input 1 (*fcType* [0 4 7 10], *fcPCs* 11, *DIChas1* 1 and *hasAscSemiToZero* 1), three properties exclusively from input 2 (*fcPCs* 1, *DIChas0* 0, *hasDescSemiToZero* 1), three common properties (*fcPCs* 5, *DIChasN1*, *hasSemiToZero* 1) and three new properties that were not present in any input space (*fcRoot* 1, *fcPCs* 8, *rootDiff* 1). Therefore, the properties of the blended space come from either input space, or are completed by logical deduction through axioms describing cadences (e.g. the pitch class 8 was added as a *fcPCs* property, functioning as a fifth of the new chord), as indicated in the parentheses next to each respective property.

(Table 2)

By blending through the amalgamation process, the generation of several blends from two input spaces is allowed. In a strict sense, a cadence that does not include a common property of the two inputs (i.e., that does not satisfy the generic space restrictions), should not be considered as their blend. However, employing the generic space in such a strict manner may potentially disallow interesting blends to be generated. For instance, the 'backdoor progression' cadence in jazz, B♭7 → Cm, would not be produced if generic space restrictions are adhered to, since it does not have a semitone movement to the tonic's root, which is a common property of both inputs (*hasSemiToZero* 1). This cadence, failing to satisfy a generic space property, as well as many others are never produced by the strict version of the presented methodology.

In this study it was regarded important to include diverse cadences that could potentially be considered as blends, even if they incorporated few of the input properties and regardless of whether these properties should normally comply to generic space restrictions. For enabling the generation of several diverse blends, the restrictions imposed by the generic space are *not* considered in the amalgamation process[3]. Additionally, acceptable cadence blends are the ones whose penultimate chord conforms to a specified dictionary of chord types (domain specific knowledge). The chord type dictionary includes some standard chords in tonal music (1-5) as well as two types that allow a wider diversity in the blends:

1.[0, 4, 7] (major),
2.[0, 3, 7] (minor),
3.[0, 4 ,7, 10] (major with minor seventh),
4.[0, 3, 7, 10] (minor with minor seventh),
5.[0, 3, 6] (diminished),
6.[0, 3, 6, 10] (half diminished) and
7.[0, 4, 6, 10] (major with minor seventh and lowered fifth).

In conceptual blending, after all blends have been generated, an evaluation process ranks them according to some optimality principles (Fauconnier & Turner, 2003); a complete description of which is outside the scope of this paper and the reader is referred to Goguen and Harrel (2010), for applications of such principles in the *Alloy* algorithm. Blending optimality in this paper is tackled through the assignment of a *salience weight* for each property that indicates the importance of specific features in cadences. Specifically, there are three grades of salience, represented as numerical weight values 1, 2 and 3, where increasing values indicate increasing salience. The weight value of each feature is assigned by hand according to basic musicological assumptions on the salience of features. Specifically for the perfect cadence, the salience weights for each properties were the following:

1.Leading note to the tonic is important (weight value 3 for the *hasAscSemiToZero* and *hasSemiToZero* properties).
2.The fact that it has the F-B tritone is relatively important (weight value of 2 for the *fcPCs* properties 5 and 11 only in the case where they are both included - if only one of them is included as *fcPCs,* it is attributed a weight value of 1).

For the Phrygian cadence, the considered important feature was the downward leading note (weight value of 3 for the *hasDescSemiToZero* property). All other properties in both input spaces were considered less salient and were thus assigned a weight value of 1.

When a blend inherits the property from an input, it is also considered to inherit its salience. Therefore, the blends that are ranked higher should incorporate as much of the input features as possible. Thus, the ranking of blends is based on the *total salience* (final row of Table 2), expressed as the sum of the feature weights it inherits from the inputs. In the case where a property is not inherited from the inputs but is generated through completion, it is assigned the default salience weight value 1.

Materials

The blending setup described above produced 84 blended cadences, all of which had some relation to both or either of the inputs. The selection of cadences for the empirical experiment was made manually, including blends from different levels of the ranking, so as to attain a maximally diverse test corpus. As already stated previously, all cadences (that were assumed to be in C minor tonality/modality) consisted of two chords, the penultimate/dominant and the final/tonic. The final chord was kept constant (C minor), thus variation between the stimuli resulted from altering the penultimate chords. Also, maximum uniformity in the formation of the chords and in voice-leading was pursued: all cadences were rendered with manual (human-made) voice-leading in four-voice harmony, with the upper voice moving upwards to the tonic (^7 - ^8, where possible), and with minimal movement in the inner voices (where attainable). Figure 3 depicts the nine cadential pairs of chords, described from a music-theoretical perspective, in the following list:

1. Perfect authentic cadence, featuring the full V^7 dominant chord that resolves to the i tonic chord without 5th, in order to achieve correct voice leading. This cadence involves functional chord progression (chords moving downwards in the circle of perfect 5ths) and strong voice-leading elements (the leading note resolving upwards to the tonic and the 7th of the dominant resolving downwards to the 3rd of the tonic, with the two active voices forming a tritone: F-B).
2. Phrygian cadence, with the ♭vii chord in first inversion resolving to the i tonic chord. This cadence is considered contrapuntal, as it is based on a pair of linear steps (the downward leading note D♭ in the lower voice resolving by a semitone to the tonic and the B♭ in the upper voice moving upwards by a whole-tone to the tonic) and involves chord root movement by an ascending 2nd (B♭ to C).
3. Tritone substitution progression, with the ♭$II^{7\flat}$ chord (German-type augmented-6th chord) leading to the tonic. The chord can also be considered an altered vii^{o7} with its lowered 3rd in the lower voice. The progression incorporates elements from the two source cadences, as it includes both leading notes (upward leading note in the upper voice and downward leading note in the lower voice), includes the tritone F-B and implies a functional dominant-to-tonic relation.
4. Backdoor progression, with the ♭VII^7 chord in first inversion, in order to achieve maximum voice-leading uniformity. This progression is mainly contrapuntal and similar to the Phrygian, but without the downward semitonal leading note, while the D in the 3rd voice can be considered a borrowed element from the perfect cadence. Also, the penultimate chord is of the same type as in the perfect cadence (major triad with minor 7th) and includes a different tritone (D-A♭), implying a functional progression in E♭ major tonality.
5. Contrapuntal-type tonal cadence, with the vii^{o6} resolving to the minor tonic. The vii^{o} is considered to have a dominant function, i.e. V^7 without its root, and it has an upward leading note in the upper voice. The removal of the downward perfect 5th in the lower voice and its substitution by a downward step (D-C) can be considered an interesting affinity with the outer voices of the Phrygian cadence.
6. Plagal-type cadence, with the $ii^{o6/5}$ progressing to the tonic. The $ii^{o6/5}$ is considered also a subdominant chord with added 6^{th} (iv^{add6}), and there is no leading note in any of the voices. The progression also features a downward perfect 4th leap in the lower voice, typical of subdominant to tonic progressions. This progression can thus be considered distantly akin to the input cadences, due to certain common chordal tones (D, F), the inclusion of a tritone (D-A♭) and similar voice-leading (D-E♭, A♭-G).
7. Minor-dominant to minor-tonic progression, utilising chords from the natural minor scale (Aeolian mode). This modal progression does not include leading tones. It can be

considered closer to the perfect cadence due to the perfect 5th relation of the chord roots, but the lack of semitonal resolution in the upper voice and of the tritone can also be considered reciprocal elements of the Phrygian cadence.

8. Altered dominant-7th chord to minor-tonic progression, with the dominant in second inversion and with its 5th lowered (French-type augmented 6th chord). This chromatic linear progression was used in the second half of the 19th-century and features two leading notes, one upward in the upper voice and one downward in the lower voice. This progression is similar to nr. 3, and can also be considered closely related to both source cadences, as it incorporates both leading notes, includes the tritone F-B and a functional dominant-to-tonic relation.
9. Half-diminished 'dominant'-7th chord to minor-tonic progression. This synthetic chord progression has not actually been used in any tonal or modal harmonic idiom, but it has been included in the experiment, since it incorporates key elements from input cadences (5th root relation, downward leading note in inner voice resolving to the tonic). Despite the perfect-5th root relation, the progression cannot be considered functional (dominant-to-tonic type), and is distantly related to cadence no. 5, since the penultimate chords are of the same type (half-diminished 7th chords).

(Figure 3)

Table 3 shows the features of the penultimate chords in the GCT format. The ranking of the 7 selected blends based on the ranking scheme described previously is illustrated in the final row of the table; the selection includes high as well as low-ranked cadences.

Experiment 1

Method

The first experiment aimed to investigate relative perception within the set of the generated cadences. A pairwise dissimilarity listening test was deemed appropriate for this purpose, as the dissimilarity matrices it produces allow Multidimensional Scaling (MDS) analysis to generate geometric configurations that represent the relationships between percepts. This in turn enables the interpretation of salient perceptual dimensions.

In the pairwise dissimilarity listening test, participants were asked to compare all pairs among the 9 cadences described in the previous section using the free magnitude estimation method. Therefore, they rated the perceptual distances of 45 pairs (same pairs included) by freely typing in a number of their choice to represent dissimilarity of each pair (i.e., an unbounded scale) with 0 indicating a same pair (for a discussion of the advantages of this approach over a bounded magnitude estimation see Zacharakis et al., 2015a). Each stimulus lasted around 4 seconds and interstimulus interval was set at 0.5 seconds. The listening test was conducted under controlled conditions in acoustically isolated listening rooms. Sound stimuli were presented through the use of a laptop computer, with an M-Audio (Fast Track Pro USB) external audio interface, and a pair of PreSonus HD7 circumaural headphones.

For the analysis of dissimilarity data between the examined cadences, this work employed a non-metric (ordinal) weighted (INDSCAL) MDS approach as offered by the SPSS

PROXSCAL (proximity scaling) algorithm (Meulman & Heiser, 2008). PROXSCAL applies an ordinal (rank order) transformation to the raw dissimilarities within each participant's responses, thus addressing the issue of the different rating scales used as a result of the free magnitude estimation approach. In turn, INDSCAL computes weights that represent the importance attributed to each perceptual dimension by each participant and then uses these weights to reconstruct an 'average' perceptual space.

Participants

Twenty listeners (age range = 18-44, mean age 24.9, 10 male) participated in the first listening experiment. Participants were students from the Department of Music Studies at the Aristotle University of Thessaloniki. All of them reported normal hearing and long term music practice (16.5 years on average, ranging from 5 to 35). All participants were naive about the purpose of the test.

Procedure

Listeners became familiar with the range of cadences under study during an initial presentation of the stimulus set (random order). This was followed by a brief training stage where listeners rated the distance between four selected pairs of cadences. For the main part of the experiment participants were allowed to listen to each pair of cadences as many times as needed prior to submitting their dissimilarity rating. The pairs were presented in random order and participants were advised to retain a consistent rating strategy throughout the experiment. In total, the listening test sessions, including instructions and breaks, lasted around thirty minutes for most of the participants.

Results

Before proceeding to the main body of the analysis for the dissimilarity data we examined the internal consistency of the dissimilarity ratings. Cronbach's alpha was .94 indicating high inter-participant reliability.

In the main body of the analysis, the dissimilarity ratings were analyzed through MDS as described above. Table 4 shows two measures-of-fit (S-Stress and D.A.F.) along with their improvement for each added dimension. A two-dimensional solution was deemed optimal for data representation as the improvement of both measures when adding a third dimension was minimal. Figure 4 shows the configuration of the cadences within this 2-D space.
(Table 4)
(Figure 4)

Simple visual inspection of figure 4 can reveal some parameters that seem to have influenced the perception of the different cadences. The 1st dimension of the space can be interpreted as 'tonal' vs. 'modal' based on the fact that all cadences featuring a leading note resolving to the tonic (upward semitone movement from B to C) cluster at the negative side while all cadences featuring an upward tone movement (B♭ to C) cluster at the positive side. The plagal cadence (No. 6) that features a duplication of the tonic is positioned almost exactly in the middle of the 1st dimension. The interpretation of the configuration along the 2nd dimension, however, is not so obvious. It could be that a combination of the inherent dissonance of the penultimate chord (as reflected by its type and voicing layout) together with its distance from the

final chord in the Tonal Pitch Space theoretical/cognitive model (Lerdahl, 2001) may explain the positions along this dimension. This notion resembles the breaking of dissonance in static 'sensory dissonance' and dynamic 'tension dissonance' suggested by Huron (2007, chapter 8). Indeed, distances in the Tonal Pitch Space (TPS) in combination with the roughness of each penultimate chord calculated by the MIR Toolbox[4] (Lartillot & Toiviainen, 2007) seem to account for the ordering of cadences along the 2nd dimension. Table 5 shows the chord distance values of each cadence according to the TPS model[5] together with the roughness of the penultimate chord and figure 5 shows the scatter plot between the 2nd perceptual dimension against a simple predictor variable (TPS distance plus roughness value). The Spearman's correlation coefficient corresponding to this scatterplot is $\rho(8) = .78$ ($p<.05$) indicating a strong relationship between this metric and the 2nd MDS dimension. It has to be noted that for a linear combination of these two components the Spearman's correlation was maximized by a mere addition.
(Table 5)
(Figure 5)

Discussion

The dissimilarity rating experiment suggests a categorical perception mode in the way cadences are perceived. This is reflected by the positioning along the 1st MDS dimension and seems to be dictated primarily by the existence of an upward semitone movement to the tonic (upward leading note) in the left-hand cadences in comparison to the lack of an upward leading note in the right-hand cadences. Two major clusters of cadences were formed based on this differentiation along with one outlier (the plagal cadence) that featured neither an upward semitone nor an upward tone to the tonic but a duplication of the tonic. The implications of categorical perception in the blending process will be discussed in the final general discussion.

The differentiation of cadences along the 2nd MDS dimension was less obvious but could be explained up to a great extent by the inherent dissonance of the penultimate chords (as expressed by the MIR Toolbox roughness calculation) together with their distances from the final chord in Lerdahl's Tonal Pitch Space. The combined influence of sensory (i.e., auditory roughness) and cognitive (i.e., Tonal Pitch Space distance) parameters has been suggested to account for the perceived tension in music (e.g., Bigand et al. 1996). The next experiment, which also includes tension among other descriptors of cadential closure, will serve to clarify whether the 2nd MDS dimension could be indeed interpreted in terms of perceived tension.

Experiment 2

Method

The second experiment was designed as complementary to the first one. Pairwise dissimilarity rating can be very useful for creating a spatial representation of the perceptual space. However, while being explicit regarding perceived similarity relationships of the objects under study, it may prove to be rather implicit when it comes to the interpretation of these relationships. Therefore, we designed a Verbal Attribute Magnitude Estimation (e.g., Kendall & Carterette, 1993a, 1993b) type of experiment whereby listeners rated the nine cadences on four descriptive scales, namely *preference*, *originality*, *tension* and *closure effect*. Originality, which is a key term for creativity evaluation (Jordanous, 2013, Hekkert et al., 2003), could also be seen as an equivalent to surprise and novelty or the opposite of expectancy, all of which have been proven very important for music perception and appreciation (Huron, 2006). The alternation between tension and relaxation is regarded as one of the key factors for musically induced emotions (Huron, 2006; Lerdahl & Krumhansl, 2007; Farbood, 2011; Lehne & Koelsch, 2015). Closure effect is a specific characteristic of musical cadences (e.g., Sears et al., 2014) as they serve the purpose of concluding phrases, sections or pieces of music. And, finally, preference measures the extent to which participants may prefer some cadences over others.

After the analysis of the acquired data, an extension of this experiment was additionally carried out. As will be explained in detail later, the ratings on originality were not very consistent across participants implying that there was a lack of a common understanding of this concept. Therefore, the same experimental protocol was repeated recruiting different participants and requesting a rating on merely two additional concepts that were regarded to be relevant to originality but at the same time more clearly defined: *expectancy* and *fit*.

The points of interest were multiple here. Firstly, we wanted to see the level of agreement between raters regarding judgements upon these scales and also to examine the potential relationships between the scales. Additionally, we sought to investigate the effect that different harmonic contexts may have on the perception of these particular cadences as expressed by the ratings. And finally, we wanted to interpret these results in the light of the perceptual cadence space generated from experiment 1 and vice versa.

Materials

Figure 6 presents the stimulus set that consisted of the nine cadences of experiment 1 positioned in two different harmonic contexts (one tonal and one modal). Each stimulus comprises a four-bar phrase, with a two-bar antecedent sub-phrase and a two-bar consequent sub-phrase. The first two-bar sub-phrase suggests the harmonic content with a four-chord progression and has two versions: the tonal version (stimuli 1-1 to 1-9) in C minor tonality and the modal version (stimuli 2-1 to 2-9) in C Phrygian mode. The second two-bar sub-phrase contains the two-chord cadential progression in slower harmonic rhythm to strengthen the effect of phrase closure, and has nine versions (the cadences of experiment 1). An attempt was made to maximize both voice-leading uniformity and harmonic idiom specification. The former condition was achieved by the use of the same sequence of melodic degrees in the upper voice for almost all stimuli: ^3 - ^2 - ^1 - ^1 - ^7 - ^1 (except stimuli 1-6 and 2-6, which do not have ^7 melodic degrees). For the fulfilment of the latter condition two four-chord progressions should be devised for each of the two versions of the first sub-phrase, containing the most characteristic elements of the two harmonic idioms. The sequence for the description of the minor tonal idiom was i - V^7 - VI - iv (emphasis on functional progressions, the dominant chord and the sharpened leading note) and the sequence for the Phrygian mode was i - ♭vii - iv - i^6 (emphasis on the lowered ^2 degree and non-functional progressions). All stimuli lasted around 9 seconds. The equipment and listening conditions were identical with experiment 1.

(Figure 6)

Participants (Group 1)

Twenty six listeners[6] (age range = 18-36, mean age = 22.7, 17 male) participated in the first listening experiment. Participants were students from the Department of Music Studies of the Aristotle University of Thessaloniki. All of them reported normal hearing and long term music practice (12.8 years on average, ranging from 6 to 25). All participants were naive about the purpose of the test.

Participants (Group 2)

Twenty five listeners (age range = 20-50, mean age = 26.7, 15 male) participated in the second listening experiment. Participants were students from the Department of Music Studies of the Aristotle University of Thessaloniki. All of them reported normal hearing and long term music practice (15 years on average, ranging from 5 to 40). All participants were naive about the purpose of the test.

Procedure

Listeners became familiar with the type of the stimuli through an initial random presentation of five examples. Then the stimuli were presented within the two different harmonic contexts. Both the order of the harmonic context and the order of the cadences within each context were randomized. Participants were allowed to listen to each stimulus as many times as needed prior to submitting their rating on all provided scales. The strengths of the attributes were represented by sliders tagged with Greek attribute names (featuring also an English translation in parenthesis) whose endpoints were labeled 'high - low' corresponding to a hidden numeric scale ranging from -10 to 10. In total, the listening test sessions, including instructions and breaks, lasted around twenty minutes for most of the participants.

Results

Before analyzing the data, we examined the internal consistency of responses for each rating scale for both harmonic contexts. Cronbach's alpha was .91 for preference, .77 for originality, .85 for tension, .94 for closure effect, .94 for expectancy and .92 for fit. These results indicate excellent inter-participant reliability for *preference, closure effect, expectancy* and *fit*. The consistency of *tension* is good, but *originality* features a significantly lower consistency. Based on this, originality will not be further examined since interpretation of its results is not considered reliable. Figure 7 presents the boxplots of each cadence for the five descriptive scales and the two harmonic contexts.

(Figure 7)

As ratings on several cadences did not pass the Shapiro-Wilk normality test ($p<0.05$) a non-parametric approach was taken for examining the effect of harmonic context on cadence description. Wilcoxon Signed-rank tests for each cadence revealed a harmonic context effect only for the expectancy ratings of the perfect (No.1) (Mdn_tonal = 10 vs. Mdn_modal = 8.8), $Z = 2.1$, $p < 0.05$, r = .29, phrygian (No.2) (Mdn_tonal = -5 vs. Mdn_modal = -1.2), $Z = -2.8$, $p < 0.05$, r = -.40 and French sixth (No.8) (Mdn_tonal = 0.07 vs. Mdn_modal = 5.3), $Z = -2.2$, $p < 0.05$, r = -.31. For all the other cadences and rating scales no effect of harmonic context was found. Furthermore, in the rating scale level, expectancy was the only scale that featured a significant effect of harmonic context indicating an overall increase in modal context (Mdn_tonal

= 0 vs. Mdn_modal = 2.4), $Z = -2.35$, $p < 0.05$, r = -0.11. Figure 8 shows the boxplot of all five rating scales aggregated for both harmonic contexts (although the overall expectancy boxplots should be viewed having in mind that this scale exhibits an effect of harmonic context). Inspection of figures 7 and 8 reveal that cadences that featured an upward leading note (No. 1, 3, 5 and 8) tended to receive higher ratings for closure effect and tension, and lower ratings for preference regardless of harmonic context. Thus, the positioning of cadences along dimension 1 of the perceptual space (Figure 4) is also reflected by the descriptive data. A Page's trend test showed a very strong trend (Page's L = 13494, z =117.28, p = 0) for increasing closure effect from the positive to the negative side of the 1st MDS dimension. This suggests that positioning of cadences along this dimension represents the perceived 'strength' of closure signified by the cadence.

(Figure 8)

The interpretation of dimension 2 is not so straightforward. A visual inspection of the boxplots for overall tension implied that tension might play a role in positioning along dimension 2. To examine this hypothesis, we performed a Page's trend test that showed a significant trend for increasing tension along the 2nd MDS dimension (Page's $L = 11749$, z = 3.20, $p << .001$). Strong trends were also present within the leading-note plus plagal cadence cluster (No. 6-5-1-8-3) and the absence of upward leading note cluster (No. 6-7-2-9-4)[7] (Page's $L = 2489$, $z = 24.80$, $p << .001$ and Page's $L = 2409$, $z = 11.50$, $p << 0$ respectively). In line with the findings of experiment one, the above also provide some evidence that dimension 2 of Figure 4 is related to perceived tension. However, the trend became even stronger (Page's $L = 12806.5$, $z = 72.33$, $p = 0$) when the ordering of cadences came from their projections on a -45^{o} line as shown in figure 9, implying that tension and closure effect (i.e., dimension 1) are not completely independent.

Table 6 shows the Spearman's correlation coefficients between the six rating scales constructed by the mean rating for each of the 18 stimuli (i.e. 9 cadences in both contexts). In agreement with the boxplots presented previously, preference features a very strong inverse correlation with closure effect, expectancy and fit (i.e., stronger closure/expectancy/fit induces less preference than weak closure/expectancy/fit). All these four variables seem to be used in essentially the same manner. Tension is the variable that conveys the highest amount of unique variance within this set being the least related to the others. Nevertheless, it shows medium correlations with closure effect (in line with what was suggested above), expectancy and fit.

(Table 6)

(Figure 9)

Discussion

Out of all descriptive qualities in the verbal attribute magnitude estimation experiment, *originality* seems to have been least understood (highest disagreement) by the listeners. This finding implies that, despite originality being a commonly agreed upon measure of creativity, it may not be a clear-cut concept within all contexts. In this particular case, it seems possible that many (but not all) listeners might have confused the concept of 'originality' (relating to novelty and inventiveness) with the concept of authenticity that relates to the root 'origin'. In this respect, we speculate that the term 'originality' might have introduced uncertainty as to whether it stands for 'novelty' or indeed 'conventionality'.

As a result of the above, we conducted an additional experiment with different participants requesting for ratings on two additional scales: *expectancy* and *fit*. These two qualities were highly agreed upon, and expectancy was the only quality that exhibited an effect of harmonic context. It could be argued that the modal context is more ‘flexible’ allowing for more possibilities; the expectancy of cadences that were unexpected in the tonal context (such as the Phrygian and the French sixth) is increased and, at the same time, the expectancy of the perfect cadence is decreased. For all other qualities no effect of harmonic context was revealed. Three factors can be taken into account as a possible explanation of this: 1) the two-bar harmonic progression that defined the tonal/modal context might have been too short to firmly establish the context, 2) the participants were more familiar with the tonal idiom, due to their prolonged exposure to classical Western music, and therefore tended to favor expectancy for tonal cadence even in stimuli with modal context, 3) the participants tended to conceive chromatic or extended cadential chords as tonal instead of modal as Renaissance modality did not include such sonorities (they were historically used only in 19th-century modality –e.g. in national musics– as exotic extensions/transformations of chromatic tonality).

The mean ratings on preference, closure effect, expectancy and fit were highly correlated showing that (in average) participants favored cadences that were less expected, i.e., had a weaker closure effect. This finding should not be generalized, however, as it might well be the case that people may tend to prefer more expected/familiar cadences within a more unexpected harmonic background. In other words, unexpectedness might be favored when introducing novelty while expectedness might be favored when resolving high uncertainty. Further experimentation is warranted to validate this hypothesis. Furthermore, closure effect, that is a direct outcome of the existence of an upward leading note (or lack thereof), seems to be the major contributor to whether two cadences will be perceived as similar, thus reflecting the categorical perception of cadences which was discussed previously. Tension is less related to the other qualities and there is indication that it may be associated with the second dimension of the perceptual space. However, tension is not completely independent from closure effect and expectancy partly confirming Huron’s (2006, chapter 8) view that these two are positively related. These results imply that, in general, the higher the expectancy (presence of an upward leading note) the stronger the tension but -according to the results of experiment 1- within each of the two groups of cadences, tension differentiations can be attributed to the inherent roughness (sensory dissonance) of the penultimate chord and the distance of the pair in the Tonal Pitch Space (tension dissonance). This is in accordance with other -complementary to Huron’s- views with regard to musically induced tension in general (Lehne & Koelsch, 2015) and tonal tension in particular (Lerdahl & Krumhansl, 2007).

A more specific look can reveal some characteristics of certain cadences. The perfect cadence gets the highest closure effect/expectancy/fit ratings and the Phrygian cadence is rated quite low for closure effect/expectancy/fit while the various products of the cadence blending system fill the space in-between. Moreover, cadences 4 (backdoor progression) and 9 (half diminished fifth) seem to get the highest preference while the perfect cadence receives the lowest preference rating in both harmonic contexts.

Therefore, despite the identified categorical perception for the cadences examined in this work, the acquired knowledge of the perceived relationships in combination with qualitative characteristics is still valuable for enhancing the creativity of the system. This information can be exploited by the cadence blending system in order to increase its capability for interaction with a human user by enabling refinement of the desired outcome. For example, when the system receives a request to produce a cadence that should be perceived relatively close (i.e., having similar closure effect) to the Phrygian but at the same time featuring higher tension, it will direct itself towards the backdoor progression (No. 4). Another example could be the request to produce a cadence that would feature a similar closure effect compared to the perfect (i.e., close in the perceptual space) but with the highest possible tension. In that case, it should direct itself towards the French sixth (No. 8). Finally, if the request is for a blend that is far away from both the perfect and the Phrygian, the plagal (No. 6) among others is a possible solution.

General Discussion

The purpose of this work was to present a case study of conceptual blending in music harmonic structures and to obtain some insight regarding the way its outcomes are being experienced by human listeners. Using two cadences (the perfect and the Phrygian) as starting points, our system produced several blends, seven of which were selected for empirical assessment. To this end, two listening experiments were conducted to shed some light on cadence perception within and out of harmonic context. Both the relative perception of these cadences and their description on (initially) four selected qualities were obtained.

From the perspective of creativity evaluation, the two input cadences (perfect and Phrygian) were positioned in the maximum distance along the 1st dimension of the perceptual space. However, no blend occupied a position that was directly in-between the original cadences, i.e., no blend was double-sided according to the results of the dissimilarity rating experiment. One could maintain that all blends (with a possible exception of the plagal cadence that is considered as an outlier) were perceived as variations of either the perfect or the Phrygian cadence.

For instance, despite the fact that some blends featured salient characteristics from both originals (such as the tritone substitution where both the leading note and the ^ ♭ 2 are present and lead to the tonic), cadence perception was categorical, based on the presence or absence of the upward leading note (the tritone substitution can be seen as a single-sided blend that preserves primarily the perfect cadence character but has embodied characteristics from the Phrygian cadence). It can be argued that the presence of the downward leading note ^ ♭ 2 in the tritone substitution cadence was overshadowed by the perceptual dominance of the upward leading note and failed to bring cadence No. 3 in the middle between No. 1 and No. 2. This does not seem to confirm our initial hypothesis that the D ♭ resolution to the tonic is equally salient to the upward leading note. In our experimental set up, this may also be due to the fact that the upper voice (that always features the upward motion in our case) is of higher perceptual salience compared to the bass (that always features the downward motion) (Thompson & Cuddy, 1989; Palmer & Holleran, 1994).

To sum up, having input spaces with a single salient property that is mutually exclusive (such as the upward semitone and upward tone to the tonic) prevents the blending system from creating balanced (i.e., double-sided) blends. This is something that may need to be taken into account in future attempts of conceptual blending of harmonic structures.

In some accordance with the dissimilarity data, the two original cadences were generally rated in the extreme values of expectancy, preference, closure effect and fit (that seem to be well represented by MDS dimension 1) and have produced seven blends that received various values in-between. Additionally, the blends received both higher and lower values of tension ratings compared to the originals. This shows that the blending system is capable of exploring away from its inputs, as Pearce and Wiggins (2001) put it, by exhibiting a variability regarding both perceptual distances and several qualitative attributes, thus highlighting its creative potential.

The hypothesis that the conceptual space represented by the perfect cadence would have higher prominence seems to be confirmed by the data. Its representative cadences induce stronger closure effect and are perceived as more appropriate endings (higher fit) regardless of context. At the same time, this increased predictability is translated into lower preference. Within the group of 'tonal' cadences, however, the French sixth (No. 8) and the tritone substitution (No. 3) seem to be more preferred, probably because of the higher amount of surprise they introduce. This is in agreement with the fact that they receive the highest positions in the system ranking in terms of blend quality and suggests that successful blending of a prominent conceptual space (in our case the perfect cadence) with a weaker one (i.e., the Phrygian cadence) has raised the preference by introducing an interesting variation. This effect was not conversely evident, however, since the Phrygian cadence was already appreciated and so were its variations through blending.

As a conclusion, this exploratory study on blending of musical cadences has demonstrated the creative potential of conceptual blending theory when applied to musical harmony. Future work will build on this approach to enable the application of the blending mechanism to more complex harmonic structures, ultimately aiming at harmonic blending between separate musical idioms. This will require the definition of harmonic concepts characterizing different idioms and their expression through a formal computational model. Assessing the ability to produce hybrid harmonic idioms or the extent to which the characteristics of an idiom are conveyed through harmony alone will be the subject of future empirical studies. This will, in turn, require the application of behavioral approaches capable of assessing longer musical stimuli where pairwise dissimilarity rating will not apply. Additionally, an advanced version of the harmonic blending system will be aimed to offer the possibility of high level descriptions of the desired harmonizations by extending the descriptive approach presented in the current work.

References

Aldwell, E., & Schachter, C. (2003). *Harmony and Voice Leading*. Belmont: Wadsworth/Thomson.

Antovic, M. (2009). Musical metaphors in Serbian and Romani children: An empirical study. *Metaphor and symbol*, *24*(3), 184-202.

Antovic, M. (2011). Musical Metaphor Revisited: Primitives, Universals and Conceptual Blending. Available online at SSRN http://ssrn.com/abstract=1763503. Accessed 14 May 2015.

Barnett, G. (2002). Tonal organization in seventeenth-century music theory. In T. Christensen (Ed.), *The Cambridge History of Western Music Theory* (pp. 407-455). Cambridge: Cambridge University Press.

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research*, *62*, 237-254.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics,* 58, 125-141.

Boden, M. A. (2004). *The creative mind: Myths and mechanisms* (2nd edition). Routledge, London, UK.
Cambouropoulos, E. (2012). A Directional Interval Class Representation of Chord Transitions. In *In Proceedings of the Joint Conference ICMPC-ESCOM 2012 (12th International Conference for Music Perception and Cognition, & 8th Conference of the European Society for the Cognitive Sciences of Music)*, ICMPC-ESCOM 2012.
Cambouropoulos, E., Katsiavalos, A. & Tsougras C. (2013). Idiom-independent harmonic pattern recognition based on a novel chord transition representation. In *Proceedings of the 3rd International Workshop on Folk Music Analysis (FMA2013)*, FMA 2013.
Cambouropoulos, E., Kaliakatsos-Papakostas, M., & Tsougras, C. (2014). 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-SMC),* Athens, Greece.
Caplin, W. (1998). *Classical Form*. New York: Oxford University Press.
Caplin, W. (2004). The Classical Cadence: Conceptions and Misconceptions. *Journal of the American Musicological Society*, 57(1), 51-118.
Collins Judd, C. (2002). Renaissance modal theory: theoretical, compositional and editorial perspectives. In T. Christensen (Ed.), *The Cambridge History of Western Music Theory* (pp. 364-406). Cambridge: Cambridge University Press.
Confalonieri, R., M. Schorlemmer, E. Plaza, M. Eppe, O. Kutz, & R. Peñaloza. (2015). Upward Refinement for Conceptual Blending in Description Logic – An ASP-based Approach and Case Study in EL++. In *International Workshop on Ontologies and Logic Programming for Query Answering, International Joint Conference on Artificial Intelligence (IJCAI) 2015*.
Cook, N. (2001). Theorizing musical meaning. *Music Theory Spectrum* 23/2: 170-195.
Eppe, M., R. Confalonieri, E. Maclean, M. Kaliakatsos-Papakostas, E. Cambouropoulos, M. Schorlemmer, M. Codescu, & K.-U. Kühnberger. (2015). Computational invention of cadences and chord progressions by conceptual chord-blending. In *International Joint Conference on Artificial Intelligence (IJCAI) 2015*.
Farbood, M. (2012). A parametric, temporal model of musical tension. *Music Perception, 29*, 387-428.
Fauconnier, G., & Turner, M. (2003). *The Way We Think: Conceptual Blending And The Mind's Hidden Complexities*. New York: Basic Books, reprint edition.
George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon.
Goguen, J. A. (2006). Mathematical Models of Cognitive Space and Time. In D. Andler, Y. Ogawa, M. Okada, and S. Watanabe (Eds.), *Reasoning and Cognition, Interdisciplinary Conference Series on Reasoning Studies*, (pp. 125-128). vo. 2., Keio University Press.
Goguen, J., & Harrell, D. F. (2010). Style: A Computational and Conceptual Blending-Based Approach. In S. Argamon, and S. Dubnov, (Eds.), *The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning* (pp. 147–170). Berlin: Springer.
Hekkert, P., Snelders, D., & Wieringen, P. C. (2003). 'Most advanced, yet acceptable': typicality and novelty as joint predictors of aesthetic preference in industrial design. *British journal of psychology*, *94*, 111-124.
Huron, D. B. (2006). Creating tension in *Sweet anticipation: Music and the psychology of expectation*. MIT press: London, UK.

Johnson, M. L., & Larson, S. (2003). " Something in the Way She Moves" - Metaphors of Musical Motion. *Metaphor and symbol*, *18*(2), 63-84.
Jordanous, A. K. (2013). Evaluating computational creativity: a standardised procedure for evaluating creative systems and its application, *Doctoral dissertation, University of Sussex*.
Kaliakatsos-Papakostas, M., Makris, D., Tsougras, C., & Cambouropoulos, E. (2016). Learning and Creating Novel Harmonies in Diverse Musical Idioms: An Adaptive Modular Melodic Harmonisation System. *Journal of Creative Music Systems* (in press).
Kendall, R. A., & Carterette, E. C. (1993a). Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck's adjectives. *Music Perception, 10*, 445-468.
Kendall, R. A., & Carterette, E. C. (1993b). Verbal attributes of simultaneous wind instrument timbres: II. Adjectives induced from Piston's Orchestration. *Music Perception, 10*, 469-502.
Koelsch, S. (2013). *Brain and Music*. Oxford: WileyBlackwell.
Lartillot, O., & Toiviainen, P. (2007). A Matlab Toolbox for Musical Feature Extraction From Audio, *International Conference on Digital Audio Effects*, Bordeaux.
Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489–508.
Lehne, M., & Koelsch, S. (2015). Towards a general psychological model of tension and suspense. *Frontiers in Psychology*, 6, 79.
Lerdahl, F. (2001). *Tonal Pitch Space*. New York: Oxford University Press.
Lerdahl, F., & Krumhansl, C. (2007). Modeling tonal tension. *Music Perception*, *24*, 329-366.
Maher, M. L., Brady, K., & Fisher, D. H. (2013). Computational models of surprise in evaluating creative design. In *Proceedings of the Fourth International Conference on Computational Creativity* (*ICCC-13*). *Sydney, Australia* (pp. 147-151).
Moore, Allan F. (2012). *Song means: Analysing and interpreting recorded popular song*. Ashgate Publishing, Ltd..
Ontañón, S., & Plaza, E. 2010. Amalgams: A formal approach for combining multiple case solutions. In *Proc. of the Int. Conf. on Case Base Reasoning*, volume 6176 of *Lecture Notes in Computer Science*, 257–271. Springer.
Ontañón, S., & E. Plaza. 2012. "On Knowledge Transfer in Case-Based Inference." In D. Agudo, Belén, and I. Watson, (Eds.) *Case-Based Reasoning Research and Development*, *Lecture Notes in Computer Science*, volume 7466. Springer Berlin Heidelberg, pp. 312–326.
Ox, J. (2014). Analogy and Conceptual Blending are Part of a Visualization Toolkit for Artists and Scientists: Introducing the Cognitive Space Transfer. *Proceedings of the IEEE VIS2014 Arts Program VISAP'14: Art + Interpretatio*n. Paris, France.
Palmer, C., & Holleran, S. (1994). Harmonic, melodic, and frequency height influences in the perception of multivoiced music. *Perception & Psychophysics*, 56, 301-312.
Pearce, M., & Wiggins, G. (2001). Towards a framework for the evaluation of machine compositions. In *Proceedings of the AISB'01 Symposium on Artificial Intelligence and Creativity in the Arts and Sciences* (pp. 22–32).
Schorlemmer, M., Smaill, A., Kühnberger, K.-U., Kutz, O., Colton, S., Cambouropoulos, E., & Pease, A. (2014). COINVENT: Towards a Computational Concept Invention Theory, In *5th International Conference on Computational Creativity (ICCC)*, Lubljana, Slovenia.
Schubert, P. (1999). *Modal Counterpoint, Renaissance Style*. New York: Oxford University Press.
Sears, D., Caplin, W. E., & McAdams, S. (2014). Perceiving the Classical Cadence. *Music Perception*, *31*, 397-417.

Sears, D. (2015). The Perception of Cadential Closure. In M. Neuwirth and P. Bergé (Eds.), *What Is a Cadence?* (pp. 253-286). Leuven: Leuven University Press.
Spitzer, M. (2003). *Metaphor and Musical Thought*. Chicago: University of Chicago Press.
Stefanou, D., & Cambouropoulos, E. (2015). Enriching the Blend: Creative Extensions to Conceptual Blending in Music. In *Proceedings of the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM 9),* Manchester, U.K.
Thompson, W. F., & Cuddy, L. L. (1989). Sensitivity to key change in chorale sequences: A comparison of single voices and four-voice harmony. *Music Perception,* 7, 151-158.
Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems* 19(7):449–458.
Zacharakis, A., Kaliakatsos-Papakostas, M., & Cambouropoulos, E. (2015b). Conceptual Blending in Music Cadences: A Formal Model and Subjective Evaluation. In M. Müller and F. Wiering (Eds.), *Proceedings of the International Society of Music Information Retrieval Conference (ISMIR 16)* (pp. 141– 147). Malaga, Spain.
Zacharakis, A., Pastiadis, K., & Reiss, J. D. (2015a). An Interlanguage Unification of Musical Timbre: Bridging Semantic, Perceptual, and Acoustic Dimensions. *Music Perception*, *32*, 394-412.
Zbikowski, L. (2002). *Conceptualizing Music*. New York: OUP.
Zbikowski, L. (2008). Metaphor and Music. In Raymond W. Gibbs Jr., (Ed.), *The Cambridge Handbook on Metaphor and Thought*. (pp. 502–524), Cambridge: Cambridge University Press.

Author Note

Correspondence concerning this article should be addressed to Asterios Zacharakis, Aristotle University of Thessaloniki, School of Fine Arts, Music Department, 54124, University Campus, Thessaloniki, Greece, E-mail: aszachar@mus.auth.gr

Footnotes

[1] http://www.coinvent-project.eu

[2] In the process of blending through amalgams, the notions of 'amalgam' and 'blend' are the same. Therefore, in the following paragraphs they are used interchangeably.

[3] In the strict version of the system, cadences 4, 6 and 7, as presented in the *Materials* section, would not have been produced by the system since their penultimate chords do not have a pitch class with semitone movement to the tonic.

[4] Using Vassilakis' algorithm.

[5] The calculations were performed with the use of the CHORD DISTANCE RULE (Lerdahl, 2001: 60). The chord distance value yielded depends on the distance between diatonic collections, on the chordal roots' distance in the circle of 5ths and on the number of non-common tones.

[6] The two groups of participants of the second experiment were different from those who took part in the first experiment.

[7] The assumed order of cadences for both groups was from negative to positive values on MDS dimension 2.

Table 1

Properties describing a cadence - an example of the perfect cadence (ending in C minor chord).

Index	Property name	Description	Value
1	*fcRoot*	Root of the penultimate chord (numeric value)	7
2	*fcType*	Type of the penultimate chord (GCT type)	[0 4 7 10]
3	*fcPCs*	Pitch classes of the penultimate chord	{7 11 2 5}
4	*rootDiff*	Root difference for the transition	5
5	*DIChas0*	Existence of common pitch class between the two chords, i.e. zero pitch interval transition (Boolean value)	1
6	*DIChas1*	Existence of upward semitone movement between any pitch classes of the two chords (Boolean value)	1
7	*DIChasN1*	Existence of downward semitone movement between any pitch classes of the two chords (Boolean value)	1
8	*hasAscSemiToZero*	Existence of ascending semitone to the tonic -	1

		leading note (Boolean value)	
9	*hasDescSemiToZero*	Existence of descending semitone to the tonic (Boolean value)	0
10	*hasSemiToZero*	Existence of upward or downward semitone movement to the tonic (Boolean value)	1

Table 2
Example of the tritone substitution cadence invention, by blending the perfect and the Phrygian cadences. Generic space elements (common properties of inputs) are shown in bold. The assignment of salience values is explained in the text.

Property's name	**Input 1 (perfect)**	**Input 2 (Phrygian)**	**Possible blend**	**salience**
fcRoot	7	10	1 (new)	1
fcType	[0 4 7 10]	[0 3 7]	[0 4 7 10] (input 1)	1
fcPCs	[7 11 2 **5**]	[10 1 **5**]	[11 1 5 8] (combination and new)	[2,1,2,1]
rootDiff	5	2	1 (new)	1
DIChas0	1	0	0 (input 2)	1
DIChas1	1	0	1 (input 1)	1
DIChasN1	**1**	**1**	**1** (both)	1
hasAscSemiToZero	1	0	1 (input 1)	3
hasDescSemiToZero	0	1	1 (input 2)	3
hasSemiToZero	**1**	**1**	**1** (both)	3
			Total salience:	21

Table 3

The penultimate cadence chords along with their respective indexes and their ranking according to their total sum of property salience weights. Cadences belonging to the ranking positions 1-29 (in bold) are the cadences produced by the system when considering the generic space restrictions in the blending process. The highest ranking is attributed to No.3 and No.8 with 21 points of total salience, followed by No. 5 and No. 9 with 17, No. 4 and No.6 with 13 and No. 7 with 12.

	input		**blends**						
index	1	2	3	4	5	6	7	8	9
pitch classes	[7 11 2 5]	[10 1 5]	[1 5 8 11]	[10 2 5 8]	[11 2 5]	[2 5 9 0]	[7 10 2]	[1 5 7 11]	[7 10 1 5]
chord type	[0 4 7 10]	[0 3 7]	[0 4 7 10]	[0 4 7 10]	[0 3 6]	[0 3 7 10]	[0 3 7]	[0 4 6 10]	[0 3 6 10]
System ranking	-	-	**1-3**	46-66	**20-29**	46-66	67-84	**1-3**	**20-29**

Table 4

Measures-of-fit and their improvement for different MDS dimensionalities.

Dimensionality	**Stress I**	**Improvement**	**D.A.F.**	**Improvement**
1D	.36	--	.87	--
2D	.20	.16	.96	.09
3D	.13	.07	.98	.02

Table 5
The Tonal Pitch Space distance for each cadence together with the roughness value of each penultimate chord.

	cadence index								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Tonal Pitch Space distance	7	9	11	9	9	6	5	8	8
roughness (Vassilakis' algorithm)	4.20	5.60	4.13	5.11	3.06	5.05	3.16	5.25	5.13

Table 6
Spearman's correlation coefficients between the rating scales.

	Preference	Tension	Closure effect	Expectancy	Fit
Preference	1.0	--	--	--	--
Tension	-.44	1.0	--	--	--
Closure effect	-.86**	.63**	1.0	--	--
Expectancy	-91**	.50*	.94**	1.0	--
Fit	-88**	.56*	.94**	.97**	1.0

** *Correlation is significant at the 0.01 level (2-tailed).*
* *Correlation is significant at the 0.05 level (2-tailed).*

Figure Captions

Figure 1. Conceptual blending between the tonal perfect cadence and a Renaissance Phrygian cadence gives rise to the tritone substitution progression / cadence (the backdoor progression can also be derived as a blend).

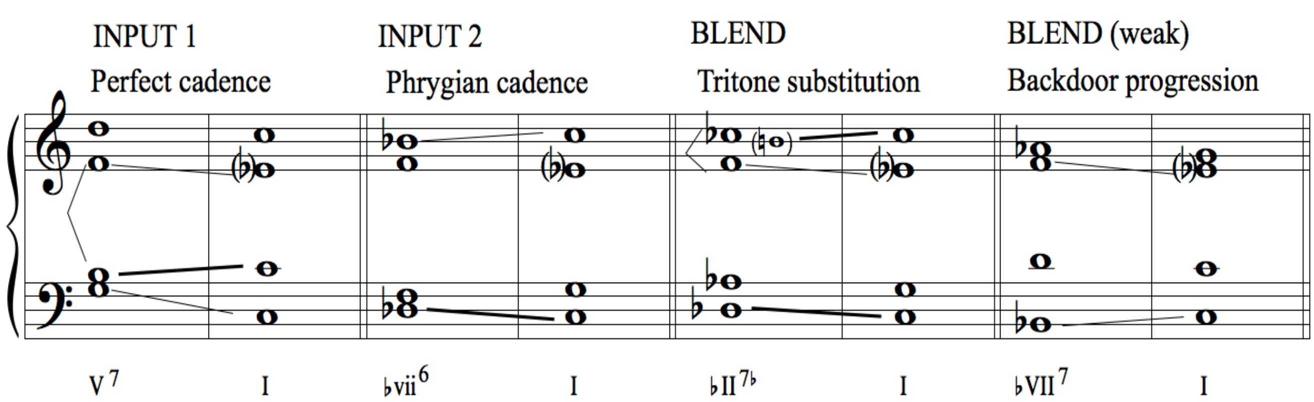
Figure 2. The conceptual blending scheme: properties describing the perfect and the renaissance Phrygian cadences are blended to create new cadences with combined properties. The generic space is computed (1) and the input spaces are successively generalized (2), while new blends are constantly created (3). Some blends might be inconsistent or evaluated poorly according to blending optimality principles or domain specific criteria.

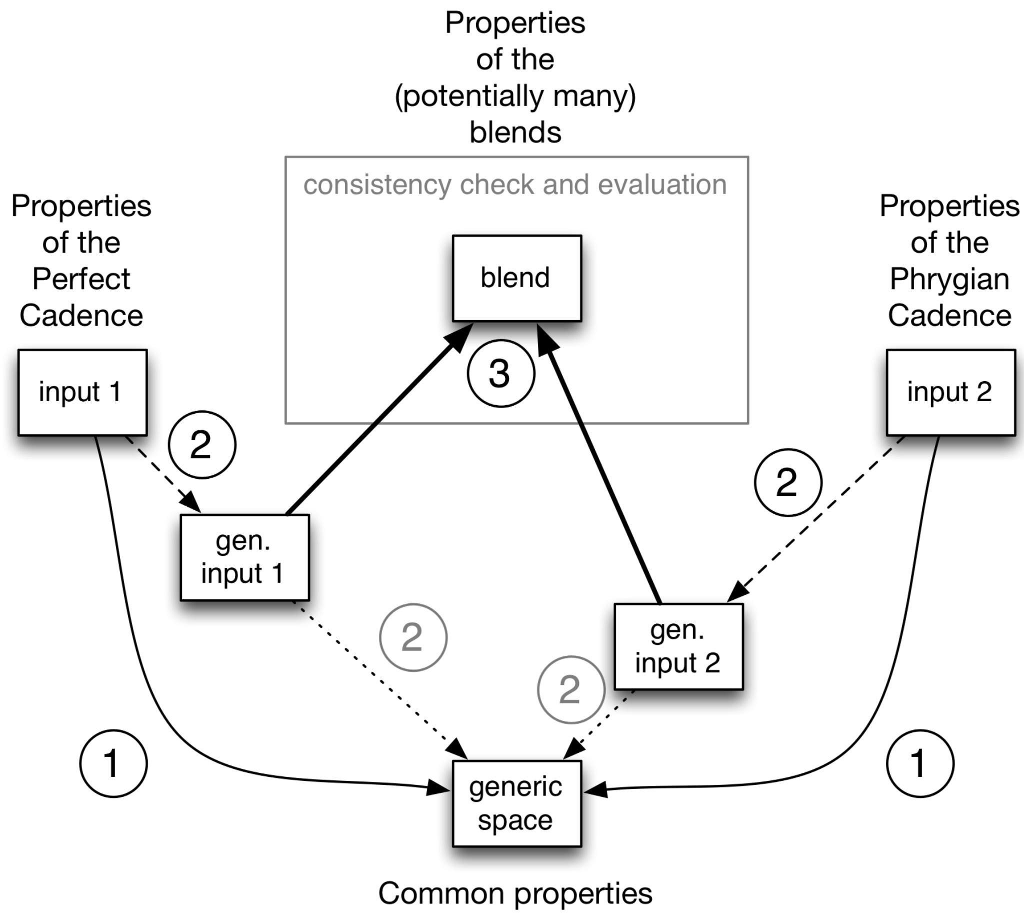
Figure 3. Score annotation of the nine cadences that constituted the stimulus set.

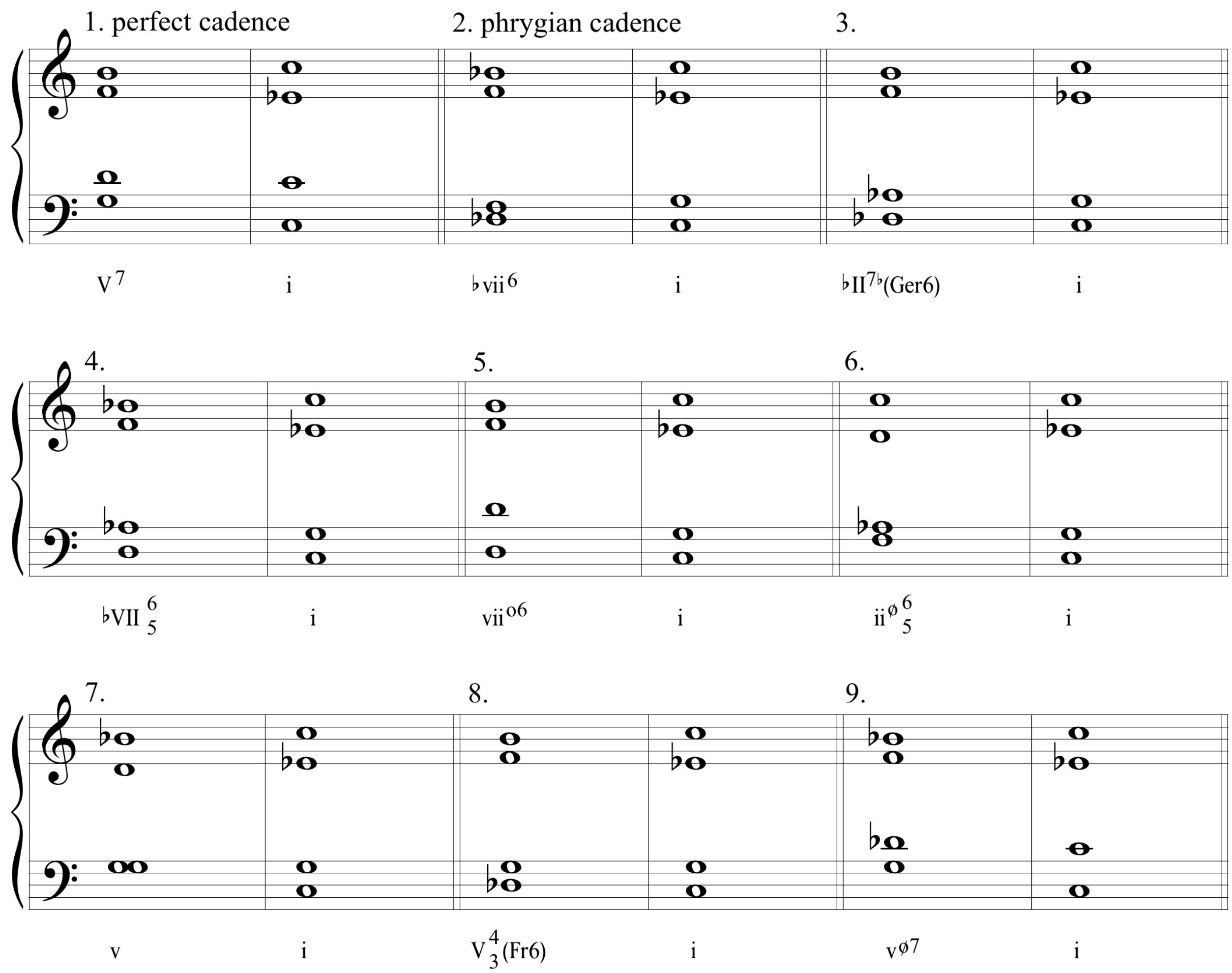
Figure 4. The two-dimensional dissimilarity perceptual space of the nine cadences. The perfect and the Phrygian cadences (No. 1 & 2) are positioned far away from each other on the 1st dimension.

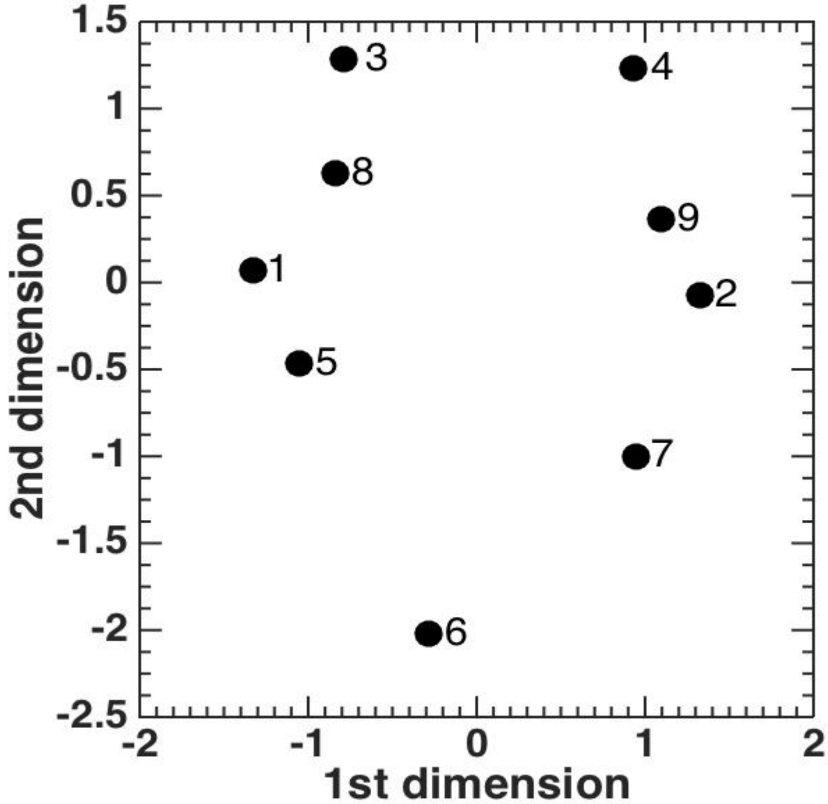
Figure 5. Scatter plot between the 2nd perceptual dimension and the simple predictor: tonal pitch space distance + roughness value.

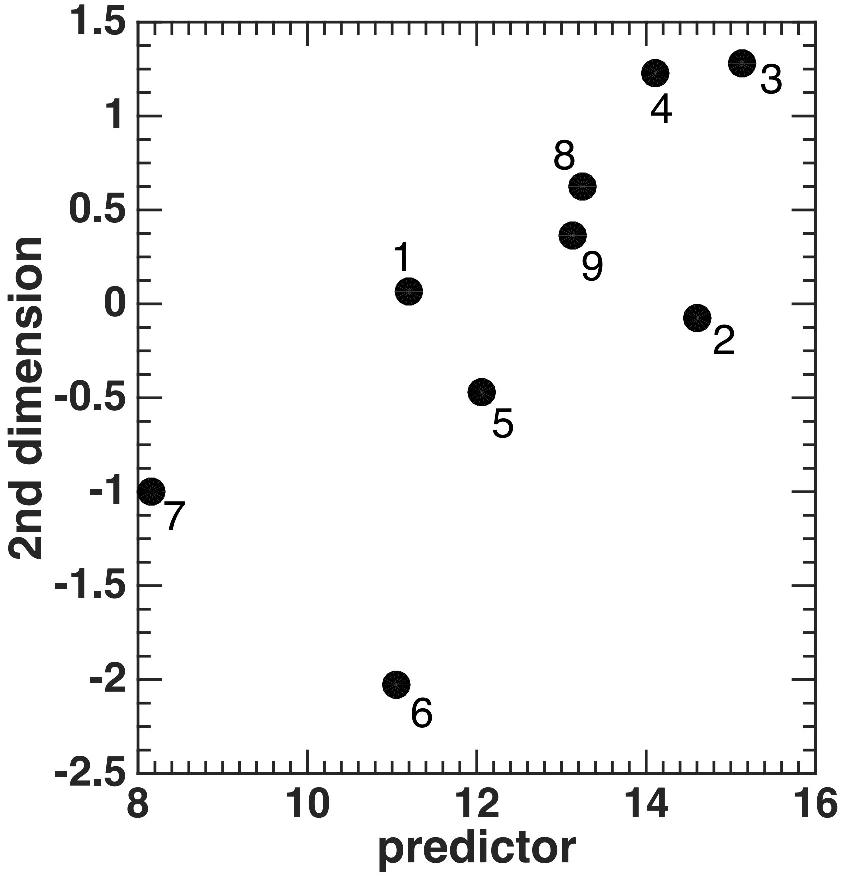
Figure 6. The score annotations of the stimulus set which consisted of the nine cadences of experiment 1 positioned in two different harmonic contexts: (a) tonal and (b) modal.
Figure 7. Boxplots of the nine cadences for the five descriptive scales and the two different harmonic contexts.
Figure 8. Boxplots of the aggregated data for the nine cadences on the five descriptive scales.
Figure 9. The perceptual cadence space with a line of -45^{o} angle. Projection on this line constitutes a good approximation of perceived tension.



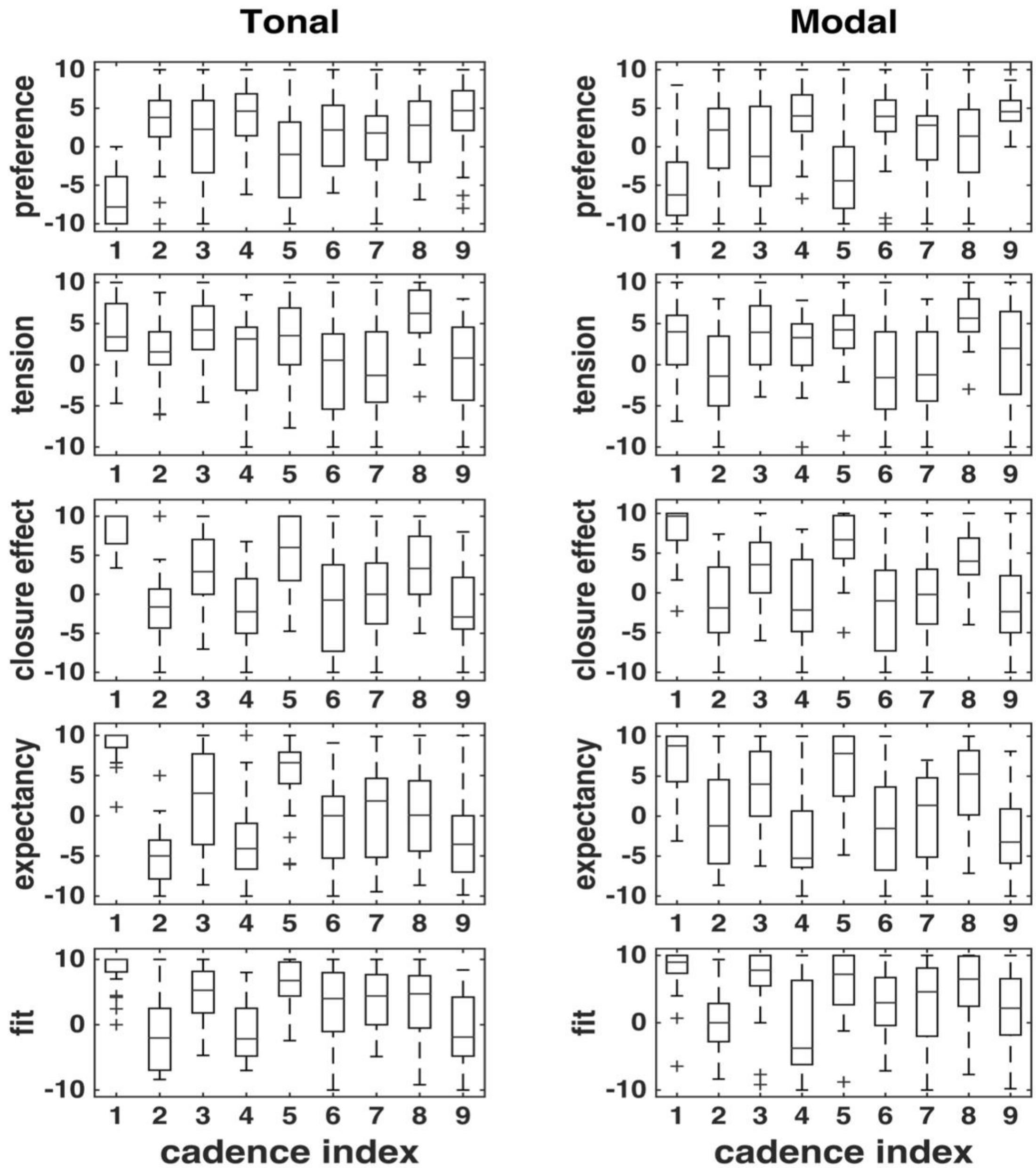

Properties
of the
(potentially many)
blends
consistency check and evaluation
blend
3
Properties
of the
Perfect
Cadence
input 1
2
gen.
input 1
2
Properties
of the
Phrygian
Cadence
input 2
2
gen.
input 2
2
1
generic
space
1
Common properties

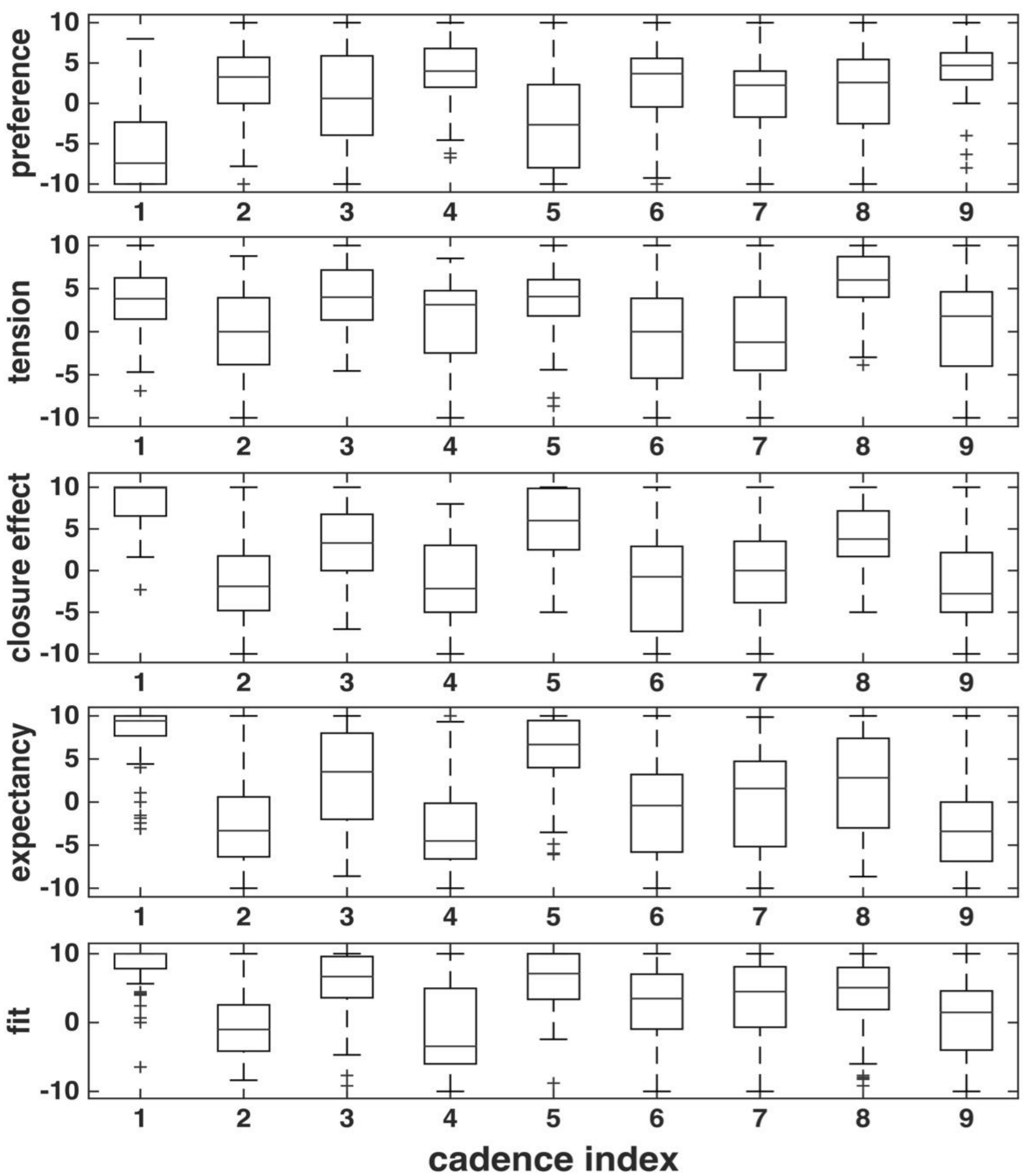

1. perfect cadence
2. phrygian cadence
3.
V7
i
♭vii6
i
♭II7♭(Ger6)
i
4.
5.
6.
♭VII 6 5
i
viio6
i
iiø 6 5
i
7.
8.
9.
v
i
V 4 3 (Fr6)
i
vø7
i


1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-2
-1
0
1
2
2nd dimension
1st dimension
1
2
3
4
5
6
7
8
9

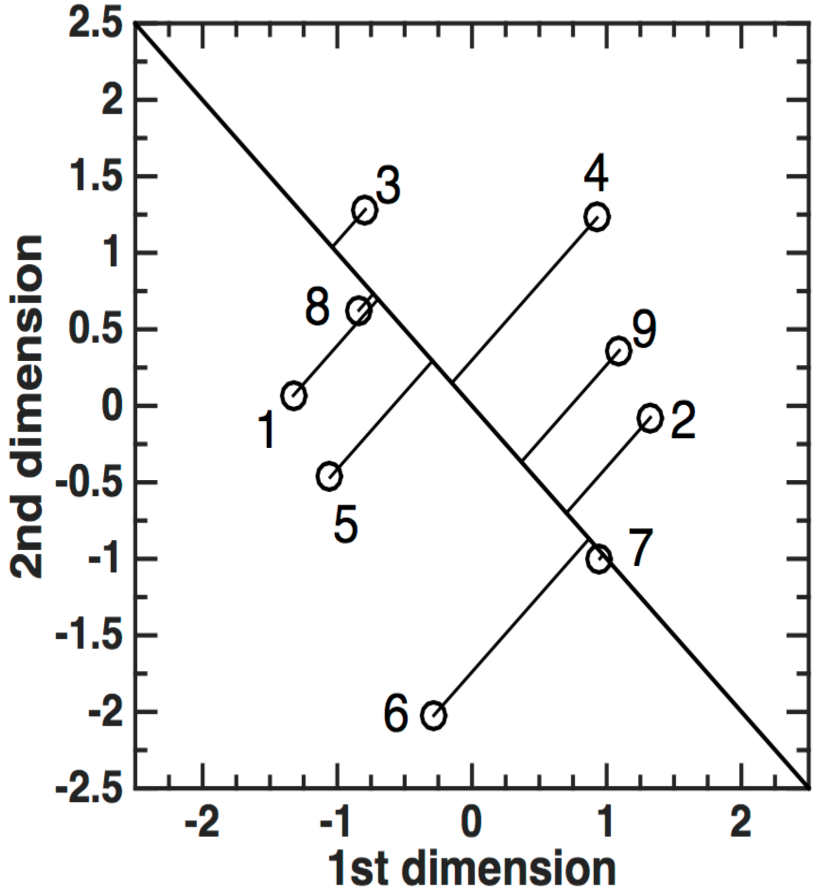

1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
8
10
12
14
16
2nd dimension
predictor
1
2
3
4
5
6
7
8
9


Cadences in minor tonal context
1-1 Tonal cadence
i V7 VI iv V7 i
1-2 Phrygian cadence
i V7 VI iv ♭vii 6 i
1-3
i V7 VI iv ♭II7♭(Ger6) i
1-4
i V7 VI iv ♭VII 6 5 i
1-5
i V7 VI iv viio6 i
1-6
i V7 VI iv iiø 6 5 i
1-7
i V7 VI iv v i
1-8
i V7 VI iv V 4 3(Fr6) i
1-9
i V7 VI iv vø7 i
Cadences in Phrygian context
2-1 Tonal cadence
i ♭vii iv i6 V7 i
2-2 Phrygian cadence
i ♭vii iv i6 ♭vii 6 i
2-3
i ♭vii iv i6 ♭II7♭(Ger6) i
2-4
i ♭vii iv i6 ♭VII 6 5 i
2-5
i ♭vii iv i6 viio6 i
2-6
i ♭vii iv i6 iiø 6 5 i
2-7
i ♭vii iv i6 v i
2-8
i ♭vii iv i6 V 4 3(Fr6) i
2-9
i ♭vii iv i6 vø7 i


Tonal
Modal
preference
tension
closure effect
expectancy
fit
10
5
0
-5
-10
1 2 3 4 5 6 7 8 9
cadence index
cadence index



preference
tension
closure effect
expectancy
fit
10
5
0
-5
-10
1
2
3
4
5
6
7
8
9
cadence index


2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
2nd dimension
-2
-1
0
1
2
1st dimension
1
2
3
4
5
6
7
8
9

Evaluating Musical Creativity: An Empirical Study on Harmonic Blending

Asterios Zacharakis, Maximos Kaliakatsos-Papakostas, Costas Tsougras & Emilios Cambouropoulos

Aristotle University of Thessaloniki, Greece

Abstract

Introduction

Creative music systems often fall into two major categories (Pearce and Wiggins, 2001): the ones whose goal is to imitate a particular musical genre or the style of a certain composer and the ones that aim to generate novel musical styles. These two approaches were also described as "empirical style modelling" and "active style synthesis" by Ames (1992). Synthesis of novel musical styles is a lot more demanding than mere modelling of certain musical characteristics as its major challenge is to transform existing rules of music making into new principles that depart from mere imitation while generating meaningful musical material. The present study constitutes an empirical evaluation of melodic harmonisations produced by a system that utilises the principles of conceptual blending theory towards active music synthesis.

Conceptual blending is a cognitive theory developed by Fauconnier and Turner (2003), which supports that even the most ordinary concepts in everyday thought are constructed through a subconscious combination of diverse, but structurally-related, mental spaces. Aspects of human creativity have been modeled using the perspective of this approach, resembling the notion of combinational creativity introduced by Boden (2009). A computational framework that extends Goguen's formal approach (Goguen 2006) has been developed in the context of the COINVENT (Concept Invention Theory) project (Schorlemmer et al., 2014). This framework aimed at a bottom-up creative approach according to which conceptual blending would be used as a creative mechanism for the generation of novel blends rather than as mere analytic instrument.

Blending in music exists in various forms, for example grafting harmonic, melodic, rhythmic or timbral elements from one musical idiom to another, thus creating novel idioms. In the context of the COINVENT framework, a melodic harmonisation assistant capable of blending between different harmonic idioms was developed (Kaliakatsos-Papakostas et al., submitted). Harmonic blending can take place through two different processes. The first –and computationally simpler- is the melody-idiom blend whereby a melody originating from a given musical idiom is harmonised based on a first-order markov model of harmonic successions (i.e., chord transition matrix) characterising another idiom. The second is the cross idiom blend whereby the chord transition matrix itself is a product of blending between two different idioms. For cross idiom blending, the definition of a -both computationally feasible and cognitively relevant- harmonic concept upon which conceptual blending would be applied was of major importance.

A precursor study on the empirical evaluation of cadence blending (Zacharakis et al., submitted) showed that applying the blending methodology to chord sequences typical of cadences (that in essence consist of a single transition between two chords) could result in the production of interesting and meaningful blends. Hence, the single harmonic transition appeared to be a promising concept for harmonic blending application. Following this, an extension of the computational approach adopted by (Zacharakis et al., submitted) to the entire chord transition matrix that characterises an idiom was implemented (a detailed description is proved in the next section). Therefore, originating from blending at the single transition level, this methodology allowed the production of extended transition matrices incorporating both transitions appearing in the two original idioms but also novel transitions that were generated through blending. It was assumed that both melody-idiom blending and cross idiom blending would reflect cognitive aspects of musical harmony. This would mean that not only harmonisations of a given melody according to a different idiom but also harmonisations based on an extended chord transition matrix would result in musical compositions featuring elements from both original idioms. One of the goals of this study was to empirically assess the extent to which these two methods of harmonic blending were cognitively relevant.

Since the developed system for harmonic blending falls into the scope of computational creativity, a comprehensive assessment of its value would require an attempt to 'measure' its creativity. A prominent debate in the field of computational creativity evaluation is whether assessment of creative systems should be based solely on the their products or also on the process through which they are generated. In a short literature overview, Jordanous (2016) presents the rationale of both approaches. While Ritchie (2007) supports that humans judge the creativity of others mainly based on what they produce, Colton (2008) argues that process may be equally important for art appreciation giving the example of conceptual art. However, while knowledge of the context in which the work of art is placed can be proven informative regarding lower-level creative processes, contextual information is (arguably) not necessarily identical to the process per se. Additionally, in their FACE/IDEA evaluating framework, Colton et al. (2011) have accompanied process assessment with audience appreciation to measure the impact of a creative act. Following this, Jordanous (2012) has suggested a system specific approach whereby the researcher is provided with a set of fourteen evaluation parameters in order to select the most appropriate ones according to the context of the assessed system. In a follow up work, Jordanous (2016) proposed that a good strategy in computational creativity evaluation may be to make an assessment based on the pair value-novelty -whose importance has been also highlighted by various researchers (Mayer, 1999; Wiggins, 2006; Ritchie, 2007; Jordanous, 2012)- under the perspective of the four Ps (i.e., Person/Producer, Process, Product, Press/Environment) (Rhodes, 1961; MacKinnon, 1970).

Creativity evaluation within the context of the present harmonic blending system is broken down into three components, the first two being value and novelty of the product. The presence of a third component has to do with the fact that the developed melodic harmonisation system had a clearly manifested goal: to utilise harmonic information from two diverse musical idioms in order to generate hybrid or novel harmonic idioms (i.e., blends). In this respect, we were interested to assess whether listeners would classify melody-idiom and cross idiom blends as blends indeed, as completely novel harmonic idioms, or as belonging to either of the originating idioms. We were also interested to examine the potential influence of melody on this process, i.e., the extent to which the implied harmony of the harmonised melody would affect idiom perception.

In order to address the above questions, as well as to evaluate the novelty and value attributed to the generated artefacts, two varied versions of a listening experiment were designed and conducted. Melodies coming from different idioms were harmonised by the system either according to the chord transition matrix of a single idiom (e.g., Bach's chorale style or Jazz) as melody-idiom blends or according to extended chord transition matrices as cross idiom blends. The task of the listeners was -with slight variations depending on the specific experimental protocol- to perform idiom classification (i.e., evaluation of idiom blending success), report their preference (i.e., attributed value) and rate the expectancy (i.e., attributed novelty) of each harmonisation.

A creative melodic harmoniser that blends harmonic spaces

The harmonic diversity in different musical styles/idioms is established by independent harmonic spaces that involve numerous concepts such as chord, root, scale hierarchy, tonality, harmonic rhythm, harmonic progression, voice-leading, implied harmony, reduction, prolongation, and so on. Conceptual blending aims to exploit the rich background~\citep{coinventGeneral2014} of concepts that is available in diverse input idioms and construct new harmonic spaces that creatively combine elements of the concepts in the input harmonic spaces. The combination of concepts from different idioms injects novelty and creativity to the melodic harmonisation process and, therefore, the existence of a rich background of diverse harmonic idioms that include formal descriptions of a variety of harmonic elements is required.

Even though the chord progressions in tonal and jazz music have been effectively modelled by models related to grammar structures ~\citep{rohrmeier_towards_2011, Koops2013, GranSteed_14}, for the purposes of blending, more musical styles need to be represented that are substantially different from the aforementioned ones. The melodic harmoniser used in this study follows a modular, hierarchical representation of harmonic structure and is able to learn from data of practically any musical idiom through statistical learning. Furthermore, this system employs conceptual blending to combine learned, diverse harmonic styles and generate new 'meaningful' harmonic idioms that can be used to harmonise given melodies. The blending methodology is based on the framework developed in the COINVENT project while it is applied on the level of chord transitions, leading to the construction of Markov transition probability matrices that blend the elements of the respective matrices in learned initial idioms. Section NEXT describes briefly the idiom-independent harmonic learning and generating methodology and the harmonic blending, while Section after_the_NEXT includes a short overview of the the harmonic blending methodology. More details for both methodologies can be found in \cite{harmoniser_JCMS} and \cite{blender_JNMR} respectively.

Statistical learning of harmonies in diverse idioms

The melodic harmoniser used for producing the material of this study is based on a statistical approach that combines different learning *modules* concerning different aspects of harmony. The utilised probabilistic algorithms allow for diverse harmonic idioms to be learned, generating harmonisations that reflect the characteristics of learned idioms in terms of *chord types*, *cadences*, *chord transitions* and *bass voice leading*. The system learns the harmonic content of an idiom through annotated training data, while it produces new harmonisations according to guidelines provided in the melody input file. The harmonic training pieces have been manually annotated by music experts in terms of the following structural aspects: (a) harmonic reduction(s) of each musical work/excerpt are created so that structural harmonic/non-harmonic notes are explicitly marked; (b) local scale/key changes are determined so that harmonic concepts relating to modulations can be learned; and (c) grouping structure is given so that cadential patterns at various hierarchic levels can be inferred.

In detail, the harmonic aspects that the system learns independently are:

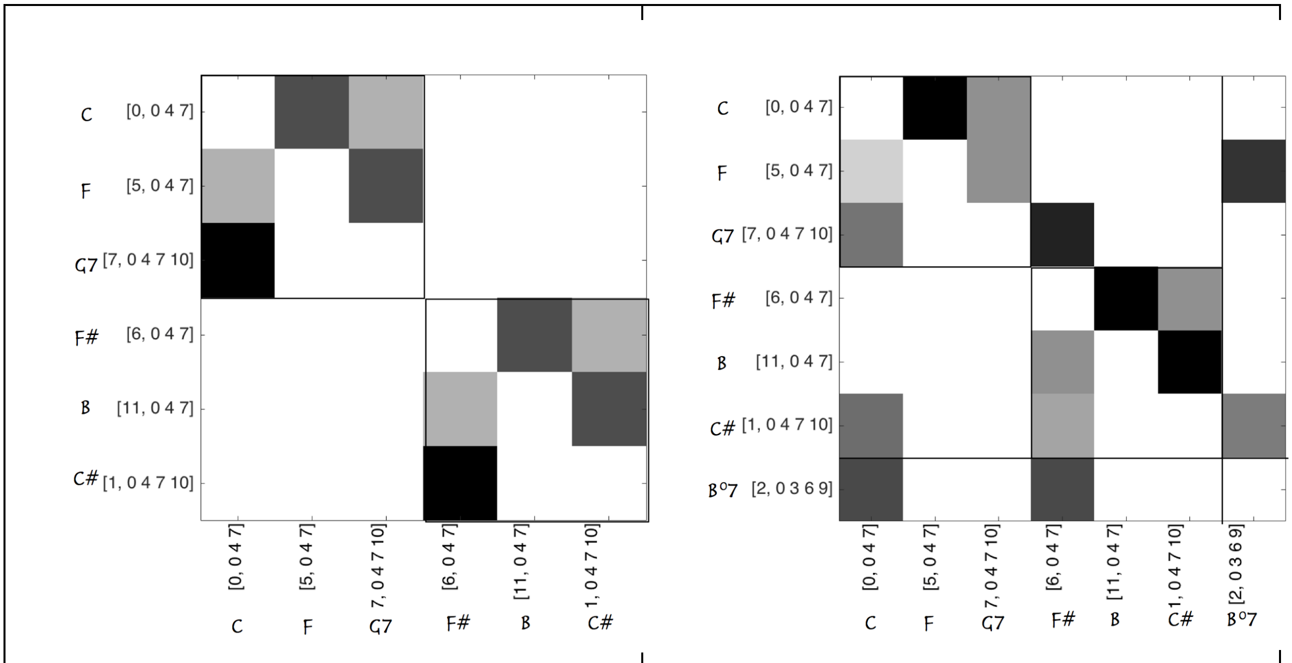
- Chords and chord types are learned in the form of the General Chord Type (GCT) \cite{cambouropoulosGCT_14,Cambouropoulos_Repr_2015}, followed by a grouping stage based on the relations between learned chord types \cite{ISMIR_gct_grouping}.
- Cadences, which are considered as the final pairs of chords in phrase endings \cite{harmoniser_JCMS}, are learned in the form of simple statistics regarding their number of occurrence at the training corpus.
- Chord progressions are learned through a model based on the hidden Markov Model \cite{rabiner_HMM}, namely the constraint HMM (cHMM) \cite{cHMM_paper}, that allows the generation of chord sequences that comply with given chord constraints (either the imposed cadences by the aforementioned module or user-defined constraints).
- Bass voice leading is learned by combining three statistical models: (a) a hidden Markov model learning the bass contour (hidden states) transitions, given the melody contour (observations), (b) distributions on the distance between the bass and the melody voice and (c) statistics regarding the inversions of chords .

After the system is trained, it can harmonise a given melody that, at this stage, is accompanied by information regarding harmonic rhythm, harmonically important notes, key and phrase structure. Learned cadences are placed at positions indicated as phrase endings in the melody input files and then chord sequences in the GCT representation are generated with the cHMM methodology. Then, the bass voice is defined by combining the bass voice related statistical models and, finally, the inner voices (between the bass and the melody) are placed according to criteria concerning attraction to a given intermediate pitch height, evenness in neighbouring note distances and movement distances of inner voices between successive chords \cite{harmoniser~JCMS}. The output of this system is a harmonic realisation with actual chord notes (not only chord symbols).

Chord transition blending for blending harmonic spaces

The melodic harmonisation system used in this work incorporates harmonic blending with an approach that focusses on chord transitions. The harmonic description of an idiom is based on the Markov matrix of GCT chord transitions, while blending is employed on the level of chord transition, i.e., transitions from one idiom are blended with ones from the other to generate blended transitions. The outcome of this process is a set of novel transitions that potentially include new chords or even new chord types, that preserve, however, some salient features of the input transitions. Transition blending can be considered as a generalisation of cadence blending as studied in \cite{Eppe_2015, Zacharakis_2015}, where different cadences were considered as pairs of chords including different penultimate chords leading to a fixed final chord; in a pair of chords forming a transition, the final chord is not fixed. However, a richer representation of transitions is used for transition blending, as discussed in this paper, in comparison to the simpler description of cadences presented in \cite{Eppe_2015, Zacharakis_2015}. The new blended harmonic idiom is a result of transition blending and several additional processes that that lead to the construction of a compound set of transitions, which comprises the transitions of the initial idioms, accompanied by a set of novel transitions and chords that provide creative and meaningful harmonic connections between the chords of the initial spaces.

To illustrate the meaningfulness of the blending approach that the utilised system follows, we use the example in Figure C-F#_FIG, taken from \cite{blender_JNMR}, where the transitions of two purely diatonic but distant harmonic spaces are blended. These spaces are assumed to incorporate only three basic chords, namely the tonic, the dominant and the subdominant, they do not overlap since the have no common chords and, as shown by the white color blocks in Figure C-F#_FIG (a), there is zero probability for transitions between chords of one space to chords of the other. In the case where the system is requested to harmonise a melody that begins in C major and then modulates to F# note, the Markov-based model of the non-blended harmonic spaces reaches a dead-end since there is no possible transition (with probability greater than 0) that leads to a chords that can harmonise this chromatic note. Aim of transition blending between the matrices representing two tonalities in this example, is to generate chord transitions that allow creative and meaningful transpositions from one tonality to the other. In general, aim of transition blending is to allow such transitions between two initial learned idioms. Therefore, transition blending creates novel transitions (as well as new chords) that enable connections between the initial harmonic spaces. The new transitions illustrated in Figure C-F#_FIG (b) are the topmost blends according to the ranking produced by a rating process that takes into account the number of common features between the blend and the input chord transitions. The considered features include common pitch classes in the first and second chords of the blend in relation to the two input transitions and ascending and descending semitone movements to the root of the final chord of each transition.



a. No Blends	(b) Topmost blends

Figure C-F#_FIG: (a) A compound harmonic space of C and F# major diatonic transition spaces that does not include blends and (b) with some of the topmost blends.

Chord transition blending can create new transitions that preserve important features of the input transitions, while the number of topmost selected transitions among the blends for further processing is determined by the user of the system with, according to a methodology described later. A larger number leads to a compound matrix that is more populated, including more transitions and chords. Similarly, the user can select the intensity of the probabilities in the transitions that result from blending, with a process that is also explained later; higher probabilities force the system to move more freely between chords of the compound space, regardless of the initial space they belong to, creating a space that radically differs for the initial harmonic spaces.

Chord transition blending and rating

The COINVENT framework for conceptual blending extends Goguen's \cite{Goguen_2016} generative model, according to which input spaces are described as *algebraic specifications* and their blended space is computed as their categorical *colimit*. According to this framework, two *input spaces* are described as sets of features, properties and relations. The description of a chord transition in this study includes properties that involve each chord independently and the entire chord transition as follows:

1. *fromPCs*: the pitch classes included in the first chord,
2. *toPCs*: the pitch classes included in the second chord,

3. *DIChas0*: Boolean value indicating whether the Directed Interval Class (DIC) vector~\citep{cambouropoulosDIC_12, cambouropoulosDIC_13} of the transition has 0 (i.e.\ that both chords have at least one common pitch class),
4. *DIChas1*: as above but for DIC value 1 (i.e., at least one ascending semitone),
5. *DIChasMinus1*: as above but for DIC value -1 (i.e., at least one descending semitone),
6. *ascSemNextRoot*: Boolean value indicating whether the first chord has a pitch class with ascending semitone relation to the pitch class of the second chord's root,
7. *descSemNextRoot*: as above but with descending semitone,
8. *semNextRoot*: as above but with either ascending or descending semitone and
9. *5thRootRelation*: Boolean value indicating whether the first chord's root note is a fifth above the root of the second. Root notes of chords are computed with the General Chord Type (GCT)~\citep{cambouropoulosGCT_14} algorithm.

After the *generic space* of two input transitions is computed, i.e. the set of their common features, an *amalgamation* process \cite{Eppe_2015, Confalonieri_2015} constructs several *amalgams* or blends. Therefore, each blend (or an amalgam) of two input spaces is a new consistent space that combines parts of the inputs and also includes (or is subsumed by) the generic space, i.e., common elements of the inputs need also to be present in any possible blend. This methodology has been successfully employed to construct new cadences as blends between the perfect cadence and the Phrygian cadence \cite{Eppe_2015, Zacharakis_2015}.

In the melodic harmoniser of this study, the notion of the generic space does not concern all transition properties, since this potentially leads to generic spaces that included many properties and allow a smaller number of ‘surprising’ blends. Properties are therefore divided into two categories: the *necessary* properties that are included in the generic space during the blending process and the *desired* that are not. If a blend does not include a *necessary* property that both both inputs have, then it is rejected. Blends that do not include *desired* properties from the inputs are not discarded, but are ranked lower, according to the rating process described later, since they do not incorporate properties that characterise the inputs. Within the context of the presented harmoniser, the necessary properties are *fromPCs* and *toPCs*, while the others are considered as desired.

The followed methodology for generating transition blends is equivalent to amalgamation, but takes advantage of the fact that a dictionary of possible GCT chord types is available in each input idiom. Therefore, in this methodology only chords within a predefined set of types (but with any possible root) can be used for generating transition blends. By employing a dictionary of acceptable chord types with size N, the number of possible chords, in terms of pitch classes, is 12 N (every chord type for each possible pitch class root) and the number of possible chord transitions is 144N^2. The use of amalgamation for exploring possible blends is thereby not necessary, since the ‘universe’ of all possible acceptable blends/transitions is not overwhelmingly large. The chord transitions that satisfy the generic space requirements are considered as acceptable blends, however, all acceptable blends need to be ranked according to some criteria that concern the *salience* of the features they include. This way, among all possible blends, the ones ranked higher are expected to capture more meaningful elements from the inputs and the overall harmonic space they are embedded in; depending on the number of possible transition blends requested (by a process that involves user selection described later), a specific number of the topmost ranked blends is selected for further processing.

In this work, the question of how meaningful a blend is, or how high is its ranking among other blends, is tackled by summing the quantified salience of the features it inherited from the input spaces. The salience value of a feature in a transition depends on the idiom that this transition is taken from and reflects how 'characteristic' or 'unique' this feature is for the examined transition in the idiom it belongs to. The numeric value attributed to the salience of a feature in a transition is computed as the fraction of one over the number of transitions that also include this feature in the examined idiom. Therefore, the greater the number of transitions in an idiom that include a feature, the less characteristic, unique, or salient this feature is for each transition that includes it. For example, the feature of having the pitch class 11 in the first chord of a transition is very salient for the transition describing the perfect cadence in tonal music, since not many transitions include a first chord with this feature. Therefore, better blends are the ones that inherit more features from the inputs that have larger salience values, a fact that is reflected by the greater sum of saliences that represent the rating value of each blend.

Using chord transition blending for constructing compound harmonic spaces

Chord transition blending is combined with learned chord transitions of the idiom-independent melodic harmonisation system discussed earlier, in order to construct new harmonic spaces that combine and extend the learned ones. Specifically, a compound chord Markov transition matrix is constructed that provides musically meaningful GCT chord transitions between learned chords of both initial idioms, using potentially new invented chords through transition blending. To this end, the 10 most usual transitions are extracted from the learned chord Markov transition tables of two initial idioms and the transition blending methodology blends these transitions, generating new, blended transitions that are afterwards imported into the compound Markov matrix. Before transition blending is employed, potential similar chords of the initial idioms are identified, enabling musically rational connections from chords of one initial idiom to chords of the other, with a process described later.

Figure FIG_COMPOUND illustrates the general form of a compound chord Markov transition matrix that extends two initial idioms, I_1 and I_2. Chords that are common in both initial idioms are considered as belonging to different sets, supposing that they potentially have different functional roles; e.g. the GCT chord [0, 0 4 7] is found both in the Bach chorales and the Jazz idioms and it is treated as a different chord in the compound matrix, i.e., there are two rows and two columns related with this chord. The sections A_{i-j} of the matrix include chord transitions that either came out of blending or by chord similarity relation (as explained later). Transitions in these sections lead directly from idiom i to idiom j. Sections B_{i-x} include transitions that lead from idiom I_i to a new chord generated with transition blending, while B_{x-i} transitions lead from a new chord to idiom I_i. Section C incorporate transitions between new chords and they are not considered in this study, since the harmoniser in this study works under the assumption that a new chord can be used only as an intermediate 'node' for transitioning from I_i to I_j.

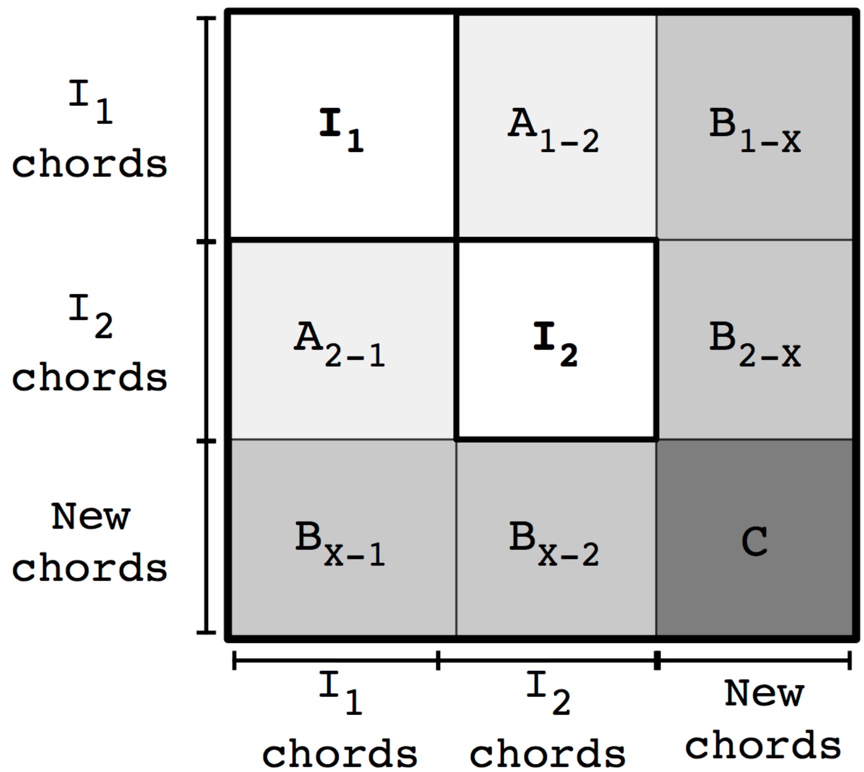


Figure FIG_COMPOUND: Graphical description of a compound matrix that includes chords and transition probabilities of both input idioms, along with new chords that were generated through transition blending.

Based on the graphical representation of a compound matrix, as depicted in Figure FIG_COMPOUND, employing transitions in section I_i leads to harmonisations that preserve character of idiom I_i, since only chords of this idiom are involved. As mentioned above, according to the assumption made for the harmonisation system of this study, a new chord created by transition blending can only be used as pivot chord that connects a chord from idiom I_i with a chord of idiom I_j. Therefore, a blended transition that includes a new second chord, e.g. c_i -> c_x, is imported in B_{i-x} if there is a transition in B_{x-j} that has c_x as a first chord, i.e. c_x -> c_j. If this requirement is not met, then c_x would potentially constitutes a 'dead-end' (terminal-only) or an 'unreachable' (beginning only) chord and, therefore, this blend is not considered for further processing.

Before blended transitions are inserted into the compound matrix, a process that identifies common or *similar* chords between the initial idioms is employed for enabling transitions between these chords with chords of the other idiom. Two chords are considered similar if they belong to the same GCT group, as defined in \cite{ISMIR15_GCTeval}, specifically if they (i) have the same root; (ii) have subset-related chord types; and (iii) both include pitch classes that are diatonic to the scale of the idiom. For example, in an idiom in C major scale, the chords [0, 0 4 7], [0, 0 4] and [0, 0 4 7 11] belong to the same GCT group (the group referring to the C major chord), while [0, 0 4 7 10] belongs to another group since it include the non-diatonic pitch class 10. According to this pre-blending process if two chords are similar in both initial idioms, then the probabilities of all transitions beginning from the chord in I_i (entire row) are copied in the row of its similar chord in A_{j-i} and all transitions leading to the chord in I_i (entire column) are copied to the column of the similar chord in A_{i-j}. The same process is employed for the similar chord in I_j.

For the application of transition blending, the 10 most common transitions in I_1 and I_2 are selected as representing the initial idioms. More transitions could be used for representing each idiom, but the number of 10 transitions is a good compromise between time efficiency and interestingness of the results. While the aim is to examine all possible blends between the 10 selected transitions in idiom 1 with the 10 selected ones of idiom 2, i.e. the results of 100 transition blending processes, to further improve time efficiency, some among the 100 possible blending processes are not executed. Specifically, transition blending process that have a generic space which subsumes (i.e., is more general than) the generic space of at least one other process are disregarded. The rationale behind this step is based on the fact that a transition blending process X, which has a generic space which is *more specific* (less general) than the one in transition blending process Y, involves input transitions that capture a larger amount of common information from their originating initial harmonic spaces. Therefore, the transition blends generated from process X are considered to be more specialised and better suited to serve as transitions between chords in the initial harmonic spaces. Transition blending processes generated by Y are a superset of the ones generated by X, but blends in the set Y-X (transitions generated by Y and not X) are more general and capture less common information of the initial harmonic spaces and, thereby, they are disregarded.

The final step is to integrate the best rated blends generated by all blending processes described above into the compound transition matrix. To this end, the topmost 100 blended transitions from each blending process are stored into a set, sorted in descending rating order. A user-defined variable in [0,1] defines what percentage of the best of these blends will be integrated in the compound transition matrix. If this variable is set to 0, then only the pre-blending connections are used in the compound matrix, while a value of 1 integrates all available blends. The probability value assigned for each transitions in this set depends on the probability values of the input transitions that produced each blend and their rating values. Specifically, the probability value of a blended transition is the product of the mean probability of the input transitions with the fraction of the rating value of this blend over the maximum rating of all produced blends. Another user-defined value allows the relative adjustment of the probability values in A_{i-j}, B_{i-X} and B_{X-i}, $i \in \{1, 2\}$. Higher values of this variable amplifies the probabilities in these regions of the matrix, promoting transitions that connect chords of I_i with chords of I_j, resulting in harmonisations that transit more often from chords of one idiom to chords of the other. Lower values promote probabilities in the I_i and I_j regions, resulting in harmonisations that incorporate larger parts of consecutive chords in a single idiom.

Method

Annotated melody files were used as inputs to the melodic harmonisation assistant for generating the stimuli that were used in the two experiments described in this section. The idioms that were mainly involved were learned from sets of Bach chorales and the Jazz pieces, while learned idioms based on sets for songs from The Beatles and pieces of Hindemith were also used. The choice to incorporate mainly the Bach chorales and the Jazz idioms in the experiments is based on the assumption that they should be known and identifiable by the participants, who were students of the Music Department of the Aristotle University of Thessaloniki. The first experiment is designed to assess the effect of blending between these idioms through perceptual tests on categorising the produced melodic harmonisations. An additional inquiry that this experiment aims to address is the effect of the implied harmony that the melody incorporates in the harmonisation process. To this end, tonal and jazz melodies were harmonised with the Bach chorales and Jazz idioms interchangeably, as well as with blended version of these two. Additional harmonisations with the learned idioms of The Beatles and Hindemith or their blended version were also used to produce material not pertaining to the Bach chorales and Jazz idioms.

Furthermore, the Bach chorales are among the most characteristics paradigms of tonal music, making them perfectly suitable for examining whether the tonal character of this idiom can be drastically altered using blending-based techniques on this tonal idiom itself. Specifically, the learned Bach chorales harmonic idiom was transposed in several keys and blending between these transposed spaces created new idioms that introduced harmonic elements that extended the tonal idiom. This effect and the extent of this diversification of the tonal spaces is assessed in the second experiment by using harmonisations of a tonal traditional melody with the idiom of the Bach chorales, a 'wrong' harmonisation with a transposed version of the Bach chorales idiom in the wrong key, transposition-related blends and an extreme harmonic blend between the Hindemith and a transposed version of The Beatles idioms.

Idiom classification experiment

Stimuli

Six blocks of stimuli were presented in the idiom blending experiment. Each of the first five consisted of a different melody that was harmonised by the system according to the tonal idiom (learned through a set of Bach Chorales), the jazz idiom (learned through a set of jazz standards) and some of their blends. In a couple of cases some harmonisations were obtained either by blending between two other idioms (Beatles and Hindemith) or according to a third idiom (Hindemith). Two of the five melodies featured tonal implied harmony (the 'Ode to Joy' theme by L. W. Beethoven's 9th Symphony and 'Ah vous dirai-je, maman' which is a French children's song used as theme in W.A. Mozart's Piano Variations K265), the other two featured Jazz implied harmony ('Summertime' by G. Gershwin and 'Someday my prince will come' by F. Churchill, soundtrack from Disney's Snow White and the Seven Dwarfs (1937)), while the last one was a Greek folk song melody ('Του Κίτσου η μάνα'). The last block consisted of a melody that was especially composed for the needs of the experiment in order to lack the third degree. Thus, a harmonisation following either a major or a minor mode was made equally possible. This melody was harmonised according to both modes and their blends. The overall number of stimuli (presented in Table 1) in all six blocks was 25.

Procedure

The experiment took place in three different sessions with simultaneous stimuli presentation through loudspeakers (M. Antovic, 2016). Each listening session featured between 10 and 20 participants. The presentation order of the blocks as well as the presentation order of the different harmonisations within each block was different for each experimental session. Listeners were provided with a questionnaire asking them to classify each stimulus in a five point likert scale between Tonal and Jazz. Apart from the three in-between positions, which implied that the stimulus could not be classified as purely Tonal or Jazz but rather as somewhere in the middle, the option of 'Other' idiom was also provided. In addition to classification, participants were asked to note their preference for each stimulus in a scale ranging from 1 to 10. As a familiarisation stage, participants first listened to one tonal and one Jazz harmonisation of the Scottish traditional melody Ye Banks and Braes and were informally asked to classify them before proceeding to the main experiment. Multiple playbacks of the stimuli were offered in the case a listener was unsure of the classification he/she should assign. Overall, including instructions, each session lasted about thirty minutes.

Participants

Forty listeners (mean age 22, age range = 18-45, 18 male) volunteered to participate in the first listening experiment. Participants were students from the Department of Music Studies at the Aristotle University of Thessaloniki. All of them reported normal hearing and long term music practice (12.8 years on average, ranging from 5 to 30). All participants were naive about the purpose of the test and especially about the fact that the creative agent under consideration was computational rather than human.

Type of chromaticism classification experiment

Stimuli

This additional experiment featured one traditional Scottish melody (Ye Banks and Braes) harmonised using the following:

1. A *tonal* idiom as learned from a set of Bach chorales (indicated by 'BC').
2. A *'wrong'* idiom obtained by transposing the Bach chorales idiom by three semitones ('BC_3').
3. A *peculiar* blend between the style of Hindemith and a transposition of The Beatles by three semitones ('BH').
4. Three *blends* between the 'correct' tonality of the Bach chorales idiom and its transposition by two, three and four semitones ('BL_2', 'BL_3' and 'BL_4' respectively).

The total number of stimuli was 7 since the tonal harmonisation was presented a twice to test consistency of the responses.

Procedure

The experiment took place in two different sessions, each one featuring 10 to 20 participants (M. Antovic, 2016). The presentation order for the different harmonisations was kept the same for all sessions as it was assumed that it would not affect the judgements due to the lack of relativity of the task. Listeners were provided with a questionnaire asking them to classify each stimulus in either one out of four categories: diatonic, chromatic, atonal and other. The smaller number of stimuli (5 vs. 25) compared to the style classification experiment (translated into shorter experimental time) allowed us to request one additional rating apart from mode class and preference for each stimulus: the degree of expectancy characterising each harmonisation. The scale of the preference and expectancy ratings ranged from 1 to 5. As a familiarisation stage, participants first listened to one tonal and one Jazz harmonisation of the Scottish traditional melody Ye Banks and Braes and were informally asked to classify them before proceeding to the main experiment. Multiple playbacks of the stimuli were offered in the case a listener was unsure of the classification he/she should assign. Overall, including instructions, each session lasted about fifteen minutes.

Participants

Thirty listeners (mean age 22.2, age range = 19-29, 13 male) volunteered to participate in the listening experiment. Participants were students from the Department of Music Studies at the Aristotle University of Thessaloniki. All of them reported normal hearing and long term music practice (12 years on average, ranging from 6 to 20). All participants were naive about the purpose of the test and especially about the fact that the creative agent under consideration was computational rather than human.

Results

Idiom and mode classification experiment

Figure 1 presents the histograms for the six categories provided for classification together with the preference ratings on the ten-point scale. Table 1 shows the excess kurtosis values (kurtosis - 3) for the preference ratings and for the ratings on the Tonal vs. Jazz classification (where the 'Other' bin was omitted). Lower kurtosis values signify the existence of more outliers (i.e., lower agreement among participants) while higher ones represent a distribution with less outliers (i.e., higher agreement among participants). It can be observed that the agreement regarding classification is generally greater when melody and harmonisation come from the same idiom (i.e., when no blending of any form takes place). Also, the mode classification (major vs. minor) distributions feature the highest kurtosis values compared to all other distributions. Preference ratings lead to somewhat flatter distributions (but not too far away from normal) for the majority of the stimuli.

Table 1. Excess kurtosis values of the classification and preference distributions for each stimulus. (Bc: Tonal harmonisation, Jz: Jazz harmonisation, Bl_L_H_M: blended harmonisation with low, medium or high blending rate respectively, Hm: Hindemith harmonisation, other: Beatles-Hindemith blend).

	Ah vous dirai-je, maman				**Greek folk song**				**Ode to joy**						**Someday my prince will come**			**Summertime**				**major-minor**			
	Bc	Jz	Bl_L	Bl_M	Bc	Jz	Bl_M	Bl_H	Hm	Bc	Jz	Bl_M	Bl_H	other	Bc	Jz	Bl_L	Bc	Jz	Bl_M	Mj	Mn	Bl_L	Bl_M	Bl_H
Classification	**4.79**	-.23	.08	.54	**-.68**	-.38	.87	-.37	-.28	**2.9**	**-.51**	**.55**	**-.94**	**-.64**	-.04	**1.7**	.67	**-.92**	-.24	-.6	**6.7**	**1.55**	**1.2**	**2.5**	**-.25**
Preference	-1.0	-.94	-1.1	-.6	.39	-.45	-.33	-.2	.28	-0.4	-.29	-.12	.47	-.51	-.17	-59	.54	.10	2.0	-.25	.09	.08	.07	-.54	.10

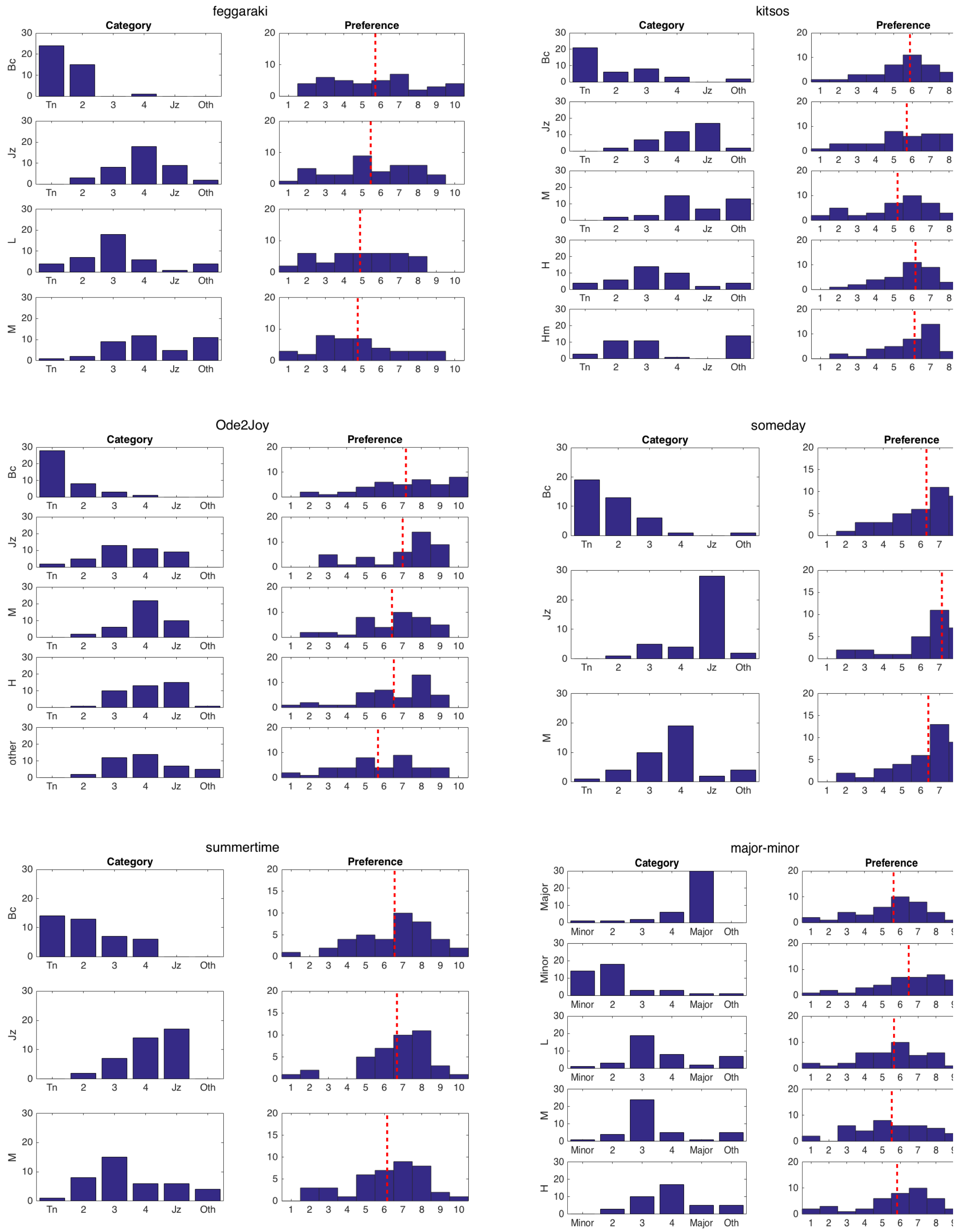


Figure 1. Histograms of the participants' responses regarding classification (left) and prefe
(right) for the different melodies and harmonisations.

Effect of harmony

Since most of the distributions failed to pass a Shapiro-Wilk test of normality, a non-parametric approach was adopted for the analysis of the data. Friedman's ANOVA tests were applied to the Tonal vs. Jazz discrete variable (once again excluding the 'Other' field) to reveal a potential effect of harmonisation on idiom classification and/or preference for the examined melodies. The results shown in table 2 indicate that while the different harmonisations indeed affected the idiom classification of every single melody, preference was affected only for *Ode to Joy* and *major-minor* at the $p<.001$ significance level.

Table 2. Friedman's ANOVA for classification (without the 'other' field) and preference.

		Ah vous dirai-je, maman	**Greek folk song**	**Ode to Joy**	**Someday my prince will come**	**Summertime**	**major-minor**
Classification	x^2	56.87	48.95	70.12	56.853	46.58	75.908
	df	3	4	4	2	2	4
	p	.000	.000	.000	.000	.000	.000
Preference	x^2	8.130	8.153	20.33	7.986	8.153	19.532
	df	3	4	4	2	2	4
	p	.043	.086	.000	.018	.079	.001

As a post-hoc analysis, Wilcoxon-Signed-Rank tests for all possible pairs of harmonisations within each melody were applied to identify the exact pairs that were classified as significantly different. Bonferroni correction (p/number of comparisons) was also applied to correct for multiple comparisons. Tables 3 to 8 present the significantly different pairs for each melody together with their effect sizes. Only four out of ten *Ode to Joy* pairs were classified as different. The harmonisation according to Bach's chorale style (i.e., the more expected tonal harmony) was classified differently compared to all other harmonisations, however no different classification was attributed to the blended (either melody-idiom or cross idiom blends) harmonisations. On the other hand, all but one *Ah vous dirai-je, maman* pairs were classified differently and the same stood for seven out of ten of the *Greek folk song* pairs, all the pairs of *Someday my prince will come* and *Summertime* and nine out of ten *major-minor* pairs. This analysis shows that, with the exception of *Ode to Joy* where the success of the blending system was partial, the harmonisations produced by the system for the rest of the melodies seemed to have generated distinguishable harmonic idioms.

Table 3. Significantly different classified pairs of Ode to Joy harmonizations identified through Wilcoxon-Signed-Rank tests with Bonferroni correction (p/10).

	Ode to Joy			
	Jz vs Bc	**M vs Bc**	**H vs Bc**	**Other vs Bc**
Z	-5.02 -	-5.5 -	-5.5 -	-5.0 -
p	<.001	<.001	<.001	<.001
Effect size	-0.56	-.61	-.62	-.59

Table 4. Significantly different classified pairs of Ah vous dirai-je, maman harmonizations identified through Wilcoxon-Signed-Rank tests with Bonferroni correction (p/6).

	Ah vous dirai-je, maman				
	Jz vs Bc	**L vs Bc**	**M vs Bc**	**L vs Jz**	**M vs L**
Z	-5.28 -	-4.88 -	-4.7 -	-4.8 +	-2.86 -
p	<.001	<.001	<.001	<.001	<.001
Effect size	-.60	-.57	-.62	-.57	-.39

Table 5. Significantly different classified pairs of the Greek folk song harmonizations identified through Wilcoxon-Signed-Rank tests with Bonferroni correction (p/10).

	Greek folk song						
	Jz vs Bc	**M vs Bc**	**H vs Bc**	**H vs Jz**	**Hm vs Jz**	**H vs M**	**Hm vs M**
Z	-5.03 -	-4.21 -	-3.96 -	-3.80 +	-4.50 +	-3.38 +	-3.71 +
p	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Effect size	-.59	-.58	-.46	-.45	-.62	-.47	-.60

Table 6. Significantly different classified pairs of Summertime harmonizations identified through Wilcoxon-Signed-Rank tests with Bonferroni correction (p/3).

	Summertime		
	Jz vs Bc	**M vs Bc**	**M vs Jz**
Z	-5.5 -	-3.8 -	-3.5 +
p	<.001	<.001	<.001
Effect size	-.61	-.45	-.40

Table 7. Significantly different classified pairs of Someday my prince will come harmonizations identified through Wilcoxon-Signed-Rank tests with Bonferroni correction (p/3).

	Someday my prince will come		
	Jz vs Bc	**M vs Bc**	**M vs Jz**
Z	-5.4 -	-4.8 -	-4.3 +
p	<.001	<.001	<.001
Effect size	-.63	-.56	-.52

Table 8. Significantly different classified pairs of major-minor harmonizations identified through Wilcoxon-Signed-Rank tests with Bonferroni correction (p/10).

	Major - Minor								
	Mn vs Mj	**L vs Mj**	**M vs Mj**	**H vs Mj**	**L vs Mn**	**M vs Mn**	**H vs Mn**	**H vs L**	**H vs M**
Z	-5.05	-4.36	-4.64	-3.73	-4.32	-4.33	-4.75	-2.80	-2.81
p	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.001	.001
Effect size	-.60	-.54	-.55	-.45	-.54	-.52	-.56	-.36	-.35

In the same manner as above, significant differences in preference as a result of harmonisation for each separate melody were further examined in a post-hoc analysis. Since the Friedman's ANOVA showed a significant effect (at the level of p <.001) only for *Ode to Joy* and *major-minor* these were the only two melodies that were tested through a *Wilcoxon-Signed-Rank* test. Table 9 shows the pairs that featured a significant difference in preference along with the effect sizes. The 'Other' harmonisation of *Ode to Joy* was significantly less prefered compared to the Bach, Jazz and medium blend and also the major harmonisation of the *major-minor* melody was less prefered compared to the minor one. Based on these results, it can be supported that, with very few exceptions, there were generally not significant differences in preference for the different harmonisations of a given melody.

Table 9. Significantly different preference for pairs of Ode to Joy and *major-minor harmonizations identified through Wilcoxon-Signed-Rank tests with Bonferroni correction (p/10).*

	Ode to Joy			major-minor
	Other vs Bc	**Other vs Jz**	**Other vs M**	**Mn vs Mj**
Z	-3.30	-3.70	-2.84	-3.72
p	.001	<.001	.004	<.001
Effect size	-.37	-.41	-.32	-.42

Effect of melody

Since we showed that the different harmonisations affected idiom classification, we were additionally interested to examine whether each of the melodies under study was affected by a different harmonisation in the same way. To this end, we considered only the harmonisations according to Bach's chorales (Tonal) and Jazz that were shared by all blocks of melodies. The assumption here was that if idiom perception was merely based on harmonic style then idiom classification would be the same for these two harmonisations regardless of the harmonised melody. A Friedman's ANOVA for Tonal and Jazz harmonisations across the five tested melodies did not confirm this hypothesis. Table 10 shows that the classification of both Tonal and Jazz harmonisations was affected by the melody itself despite not considering the 'Other' field. More specifically, the post-hoc *Wilcoxon-Signed-Rank test*s presented in table 11 showed that the tonal harmonisation of the jazz melody of *Summertime* was classified as less tonal than the ones of the tonal melodies of *Ode to Joy* and *Ah vous dirai-je, maman*. Additionally, the Jazz harmonisation of the jazz melody of *Someday my prince will come* was classified as more jazz than the jazz harmonisation of the tonals *Ode to Joy* and Ah vous dirai-je, maman. Finally, even the jazz harmonisation of the modal melody of *the Greek folk song* was classified as more jazz than the jazz harmonisation of the tonal *Ode to Joy*.

Table 10. Friedman's ANOVA (without the 'other' field) for comparison between the same harmonisation style across the five melodies.

	Tonal	Jazz
x^2	14.92	29.52
df	4	4
p	.005	<.001

Table 11. Significantly different pairs identified through Wilcoxon-Signed-Rank tests with Bonferroni correction (p/10) .

	Tonal			Jazz	
	Ode to Joy vs. Summertime	**Ah vous dirai-je, maman vs. Summertime**	**Ode to Joy vs. Greek folk song**	**Ode to Joy vs. Someday my prince will come**	**Someday my prince will come vs. Ah vous dirai-je, maman**
Z	-3.3	-3.09	-3.42	-4.29	-3.29
p	.001	.002	.001	.000	.001
Effect size	-.37	.35	.39	-.49	-.39

Type of chromaticism classification experiment

Figure 2 presents the histograms for the four categories provided for harmonic style classification together with the preference and expectancy ratings on the five-point scale. Similarly to the previous analysis, table 12 presents the excess kurtosis values (kurtosis - 3) for all three distributions (treating the style classification as a discrete rather than a categorical variable for depiction purposes) in order to get a quantification of the outlying values. The results of a Friedman's ANOVA on the preference and expectancy distributions (shown in table 13) revealed that there was an effect of harmonisation on both properties. Tables 14 and 15 present the significantly different pairs of harmonisations as resulted from a post-hoc analysis (i.e., Wilcoxon Signed-Rank tests).

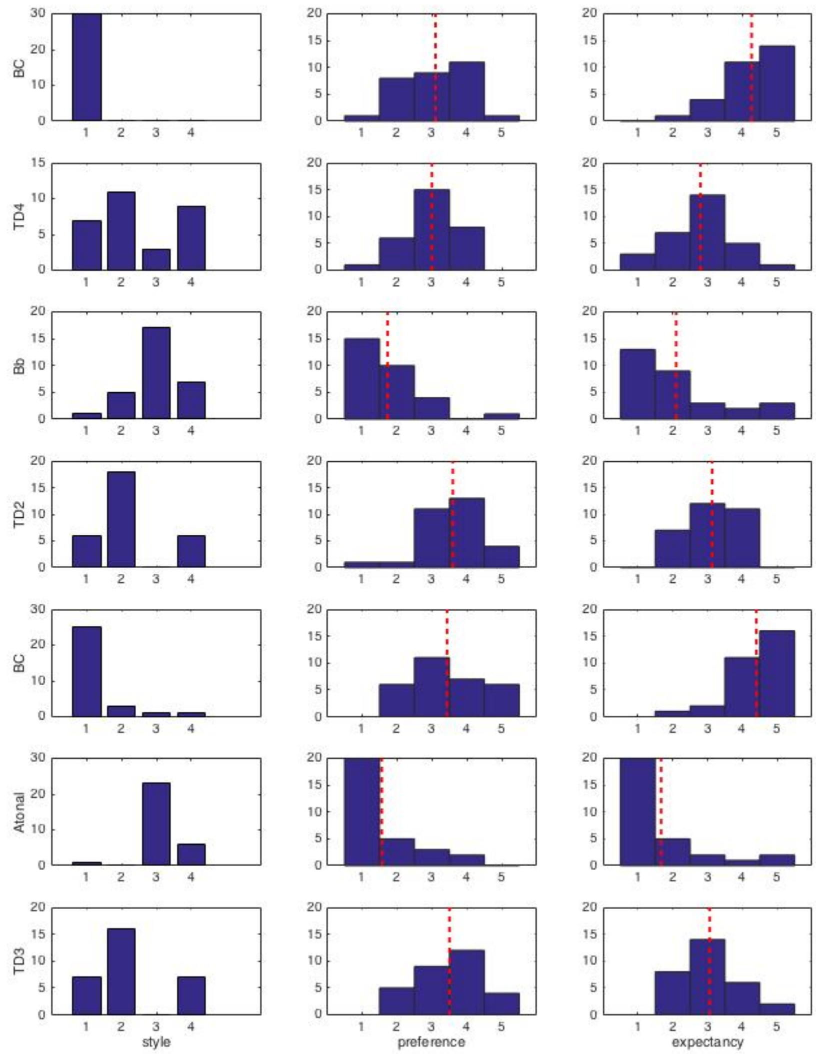


Figure 2. Histograms of the responses regarding style classification (left), preference (middle) and expectancy (right) for the different harmonisations of Ye Banks and Braes. (1: Diatonic, 2: Chromatic, 3: Atonal, 4: Other)

Table 12. Kurtosis values for mode categorization, preference and expectancy responses for the different harmonisations of Ye Banks and Braes.

	Excess Kurtosis

	BC_1	TD4	Bb	TD2	BC_2	AT	TD3
Style Classification	-	-1.44	.55	-.09	8.88	6.25	-.63
Preference	-.077	-.03	3.52	1.26	-1.07	1.38	-.77
Expectancy	.35	.05	.26	-1.26	1.87	2.85	-.14

According to figure 2 and table 12 the BC harmonisation was unanimously classified as diatonic in the first presentation and the same with very high agreement in the second exposure where it was mistaken for chromatic, atonal or other from only four out of thirty participants. In addition, it was rated as the most expected harmonisation and it was more preferred compared to Bb and Atonal. The Bb and the Atonal harmonisations were classified as Atonal with high agreement and were attributed the highest unexpectancy (with an exception between B ♭ and TD4 where the difference was not significant) and the least preference among the rest harmonisations. The TD2, TD3 and TD4 were mostly rated as chromatic but with some percentage of participants rating them as diatonic or even as other. The TD4 was the harmonisation that featured the least agreement among participants for style classification. These three harmonisations were rated as less expected than BC but most expected than Bb and Atonal. The difference of expectancy among them was insignificant. Regarding preference, TD4 was significantly less preferred than TD3 but no significant difference was found among any of these blends and BC.

Table 13. Friedman's ANOVA for expectancy and preference between the various harmonisations of Ye Banks and Braes (N=30).

	Expectancy	Preference
x^2	118.05	85.40
df	6	6
p	<.001	<.001

Table 14. Significant difference for expectancy from post-hoc analysis (Wilcoxon signed rank test) of all pairs (Bonferroni correction p/21).

	BC_1	TD4	Bb	TD2	BC_2	AT	TD3
BC_1	-						
TD4	X	-					
Bb	X		-				
TD2	X		X	-			

BC_2		X	X	X	-		
AT	X	X		X	X	-	
TD3	X		X		X	X	-

Table 15. Significant difference for preference from post-hoc analysis (Wilcoxon signed rank test) of all pairs (Bonferroni correction p/21).

	BC_1	**TD4**	**Bb**	**TD2**	**BC_2**	**AT**	**TD3**
BC_1	-						
TD4		-					
Bb	X	X	-				
TD2			X	-			
BC_2			X		-		
AT	X	X		X	X	-	
TD3		X	X			X	-

Discussion

As discussed in the introduction, a creative agent may be evaluated both in terms of the processes it incorporates and in terms of the artefacts it generates. This piece of research has focused on empirical evaluation of a number of computer-generated melodic harmonisations (i.e., products of the harmonic blending system). The most direct way to evaluate creativity would be to ask the participants' opinion on the extent of creativity demonstrated by each product (in our case harmonisation). However, since this is essentially equivalent to judging the creativity of the producer, it might introduce some biases by causing subconscious assumptions regarding its identity (e.g., human or computational). Besides, this computational system had a well-manifested target, which was the creation of hybrid harmonic idioms through harmonic blending. In this sense, there was a clear criterion for measuring success: idiom classification. By additionally requesting judgements on preference and expectancy it was also aimed to assess the value and novelty of the harmonisations.

What the empirical experiments have shown was that the different harmonisations produced by the system have indeed influenced idiom perception. The harmonisations according to purely tonal or jazz transition matrices were mostly classified as belonging to the originating harmonic style. On the contrary, harmonic blends between tonal and jazz harmony were mostly perceived as belonging to either a hybrid jazz/tonal idiom or to an unidentifiable ‘other’ idiom. The ‘other’ field was also more prominent in the two cases where the harmonisation style was not one of the tonal, jazz or their blends (i.e., Hindemith and a blend between Beatles and Hindemith). This shows that participants were -up to some extent- able to discriminate between a blend and a harmonisation that was based on a totally different idiom. At the same time, inspection of the kurtosis values of the classification distribution showed that the harmonisation of a melody by an idiom other than its original (either blended or not) generally lowers agreement among participants regarding the perceived musical style. That is, blending (either melody-idiom or cross-idiom) seems to introduce some uncertainty for idiom identification.

The identified effect of harmony on idiom perception was also melody dependent, meaning that idiom classification was decided by consideration of both harmonic and melodic characteristics. This was tested on harmonisations based on purely tonal or jazz style so as to have the same reference for all melodies. This finding implies that even a melody-idiom blend (i.e., harmonisation of a melody according to an idiom other than its original) may constitute a perceivable type of blend. In addition, preference ratings were significantly lower only for the Ode to Joy ‘other’ harmonisation in comparison to the other versions.

The mode classification (major-minor) experiment showed that people were able to classify the major mode very successfully. The slight confusion regarding minor mode classification is due to the fact that the short excerpt was concluded with a Picardy third that is common in Bach’s minor mode chorales. This has expectedly shaped the judgments of the listeners slightly towards major mode. Other than that, the blends were mostly classified as in-between major and minor mode or rated as belonging to the ‘other’ mode. In addition, the preference ratings were significantly different only between the purely minor and major harmonisations (with a preference for minor).

Finally, the type of chromaticism clasification experiment showed that judgments on compositional style were quite consistent except for TD4 harmonisation, ratings of which were almost equally distributed between tonal, chromatic and other. The tonal harmonisation according to Bach’s chorale style received a unanimous tonal rating in the first presentation with very slight differentiation in the second one. The B♭ harmonisation was mostly judged to be atonal as was the Atonal one. The TD2 was generally considered to be chromatic as was TD3 but with slightly less agreement (small leakage towards tonal and other in both cases). B♭ and Atonal received significantly lower preference ratings accompanied by significantly lower expectation ratings; a confirmation of excess novelty not being highly appreciated (Margulis and Beatty, 2008) and of the fact that the harmonisations deliberately produced to sound ‘wrong’ were identified as such. On the other hand, the BC harmonisation received significantly higher ratings for expectancy compared to all the rest but not any significant difference when preference was concerned (with the exception of B♭ and Atonal). This means that the blending system produced some novel harmonisations (i.e., TD4, TD2 and TD3) that were both recognised as being different stylistically and were equally appreciated in comparison to a conventional tonal harmonisation.

Overall, based on the above it can be supported that the harmonic blending system has indeed succeeded in producing perceivable blends -both across idioms, modes and types of chromaticism- that were equally preferred compared to non-blends. Here it is worth noting that, even though it is assumed that successful style classification provides a self-rewarding experience and therefore should positively influence aesthetic judgements (Leder et al., 2004), it has not affected preference ratings in a consistent way. The Ode to Joy 'other' harmonisation has indeed received significantly lower preference ratings combined with increased 'other' classification but this was not the case for the Greek folk song M and Hm harmonisations or the TD4 harmonisation. Of course, the fact that some participants reported the 'other' field does not necessarily mean that they had difficulty identifying some style that did not fall into the provided categories. However, the most probable explanation for this may be that, when music appreciation is concerned, successful style identification is an elementary level that can only be rewarding once the task is not completely trivial. The complete inability to perform style classification might indeed lead to negative aesthetic judgements but a situation in-between triviality and absolute unpredictability whereby despite some cues style cannot be classified with indisputable certainty may be also appear compelling. This study has not provided with evidence to either support or reject this hypothesis as the differences in preference were minimal. It has to be noted, however, that alterations between stimuli in this experiment concerned harmony alone, which is only one out of various parameters that define musical style.

Future work will aim to quantify the potential influence of the harmonisation assistant on human creative processes. Such an evaluation would not be possible through the passive listening protocol adopted in this work. Some active interaction with the system would be required instead. In this direction, a pilot study that aimed to assess the enhancement of human creativity through interaction with the system was conducted. The objective was to examine whether (and in what way) the harmonisation of a certain melody by humans would be affected as a result of participants' exposure to a number of different harmonisations of the same melody produced by the system. This would, in turn, require a reliable measure of harmonic dissimilarity that could be used as a metric of harmonic divergence (i.e., inspiration offered by the products). This piece of research is still in progress and results will be reported in future work. The obvious next step would be to allow people to make use of the system without any type of mediation to evaluate both user experience and artefact production in such a scenario.

References

Antovic, M., Stamenkovic, D., & Figar, V. (2016). Association of Meaning in Program Music: On Denotation, Inherence, and Onomatopoeia. *Music Perception, 32,* .

Boden, M. A. (2004). *The creative mind: Myths and mechanisms* (2nd edition). Routledge, London, UK.

Colton, S. (2008). Creativity versus the perception of creativity in computational systems. Proceedings of AAAI symposium on creative systems (pp. 14–20), Stanford, California, USA.

Colton, S., Charnley J., & Pease A. (2011). Computational creativity theory: The FACE and IDEA descriptive models. *Proceedings of the 2nd international conference on computational creativity* (pp. 90–95). Mexico City, Mexico.

Fauconnier, G., & Turner, M. (2003). *The Way We Think: Conceptual Blending And The Mind's Hidden Complexities*. New York: Basic Books, reprint edition.

Goguen, J. A. (2006). Mathematical Models of Cognitive Space and Time. In D. Andler, Y. Ogawa, M. Okada, and S. Watanabe (Eds.), *Reasoning and Cognition, Interdisciplinary Conference Series on Reasoning Studies*, (pp. 125-128). vo. 2., Keio University Press.

Jordanous, A. (2012). Evaluating computational creativity: A standardised procedure for evaluating creative systems and its application *(Ph.D. dissertation). University of Sussex*, Brighton, UK.

Jordanous, A. (2016) Four PPPPerspectives on computational creativity in theory and in practice, *Connection Science*, 28:2, 194-216, DOI: 10.1080/09540091.2016.1151860.

Kaliakatsos-Papakostas, M., Queiroz, M., Tsougras, C., & Cambouropoulos, E. Conceptual blending of harmonic spaces for creative melodic harmonisation. Journal of New Music Research, (Submitted – Under review)).

MacKinnon, D. W. (1970). Creativity: A multi-faceted phenomenon. In J. D. Roslansky (Ed.), *Creativity: A discussion at the nobel conference* (pp. 17–32). Amsterdam: North-Holland Publishing Company.

Margulis, E. H., & Beatty, A. P. (2008). Musical style, psychoaesthetics, and prospects for entropy as an analytic tool. *Computer Music Journal*, 32(4), 64-78.

Mayer, R. E. (1999). Fifty Years of Creativity Research. In R. J. Sternberg (Ed.), *Handbook of creativity*, Chapter 22 (pp. 449–460). Cambridge, UK: Cambridge University Press.

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British journal of psychology*, *95*(4), 489-508.

Pearce, M., & Wiggins, G. (2001). Towards a framework for the evaluation of machine compositions. In *Proceedings of the AISB'01 Symposium on Artificial Intelligence and Creativity in the Arts and Sciences* (pp. 22-32).

Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17, 67–99.

Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan*, 42(7), 305–310.

Schorlemmer, M., Smaill, A., Kühnberger, K.-U., Kutz, O., Colton, S., Cambouropoulos, E., & Pease, A. (2014). COINVENT: Towards a Computational Concept Invention Theory, In *5th International Conference on Computational Creativity (ICCC)*, Lubljana, Slovenia.

Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7):449–458.

Zacharakis, A., Kaliakatsos-Papakostas, M., Tsougras, C., & Cambouropoulos, E., (in press). Creating musical cadences via conceptual blending: Empirical evaluation and enhancement of a formal model. *Music Perception*, (submitted, under review)