Fuzzy Relational Music Perception: Resemblance and Implication in the Individuation and Assembly of Concepts

Trevor Rawbone

Jeddah International College, Saudi Arabia
t.rawbone@jicollege.edu.sa

Abstract

Fuzzy relational music perception concerns the representation of congruent connections between musical features as fuzzy relations used to individuate and assemble concepts and conceptual hierarchies. This article presents two universal fuzzy domains of discourse, harmony $\boldsymbol{H}$ and grouping $\boldsymbol{G}$, which partition sets using triangular norms (t-norms) based on *generalised harmonic root support* and *generalised time regularity*, respectively. Fuzzy relations between the sets of the domains are formed in the innate fuzzy neural architecture of a dedicated music faculty. Fuzzy relations are shown to be necessary representations for interconnection between the domains to individuate and assemble concepts. Concepts are individuated and assembled by virtue of fuzzy set *resemblance* relations between domains, or fuzzy logical *implication* relations in one or both domains through time. Fuzzy resemblance relations comprise the properties of weak reflexivity, weak symmetry and antitransitivity in a $\boldsymbol{H} \times \boldsymbol{G}$ Cartesian product space. Fuzzy implication relations involve fuzzy overlap (or continuation) of elements, calculated using a t-norm operator (min operator), in one or both domains of the product space. Supplementary theory is incorporated to explain polyphonic structure, involving pluralistic superimposition of independent fuzzy relational hierarchies. Broadly, fuzzy relational music perception is a rationalistic model that builds on generative theories and associative–statistical and connectionist approaches by providing a compact and coherent process for determining interaction across musical parameters.

Keywords: Music perception, fuzzy set theory, fuzzy logic, resemblance relation, implication relation, multiparametric congruence, concept individuation.

1 Introduction: From Generative Theories to Fuzzy Relational Music Perception

This article proposes a novel model of music perception, fuzzy relational music perception (hereafter, FRMP), which represents connections between musical features as fuzzy relations that are used to individuate and assemble concepts. This section provides theoretical foundations in review of generative music theory and fuzzy music research. The music-theoretical notion of multiparametric congruence (MC), or simply "congruence", is the relational principle that the model aims to explain and is defined as the correspondence between similarly stable elements of two or more parametric features. There is MC between such stable features as consonant chords and phrase grouping onsets, or between phrase onsets and higher levels of metrical structure. Generative music theories and generative computational implementations implicitly and explicitly incorporate congruent relations. However, it is determined that generative rule systems often result in interactional incoherence between parametric features. Fuzzy logic approaches to music generation are reviewed to explore techniques that can systematise and formalise congruent relations. The overview finds that current work in fuzzy music generation does not employ sets or relations between relevant levels of musical hierarchies to enable modelling of the individuation of chords, grouping, tonality, metrical structure and other fundamental concepts of music perception. FRMP addresses this by showing that perception of fuzzy relations in music concerns two central domains, harmony $\boldsymbol{H}$ and grouping $\boldsymbol{G}$, from which all parametric features (/concepts) are generated. The model centres on two types of fuzzy relations to describe MC: *resemblance* and *implication*, which are represented in an $\boldsymbol{H} \times \boldsymbol{G}$ Cartesian product space and incorporated by perception as necessary and sufficient conditions to individuate and assemble universal music concepts.

1.1 Multiparametric Congruence in Generative Music Theory

Generative theories and generative computational implementations incorporate congruence to connect different parametric features of music (e.g., Schenker, 1935; Cooper and Meyer, 1960; Lerdahl and Jackendoff, 1983; Lerdahl, 2001; London, 2004; Hamanaka et al., 2006; Marsden, 2010; Marsden et al., 2018). The core of the generative music research program, set out in *A Generative Theory of Tonal Music* (GTTM) (Lerdahl & Jackendoff, 1983) and *Tonal Pitch Space* (TPS) (Lerdahl, 2001), is a comprehensive conceptual and meta-conceptual system of formal and semi-formal rules that determines congruent connections between conceptual hierarchies of musical structure. The basic formal rules of the GTTM-TPS program denote crisp and idealised syntactical structures, termed well-formedness rules (based on Chomsky, 1957), which set out the fundamental architecture or structural possibilities of musical representations. To enable interaction between these musical parameters, semiformal rules, termed preference rules (/constraints), prescribe congruent preferences between the well-formedness rules, in and across four components of the theory – grouping, metre, time-span reduction and prolongational reduction. For example, there is a preference rule for stable metrical and consonant harmonic phenomena to coincide and a preference rule for stable grouping elements and harmonic change to be coupled, among others.

In a departure from the generative program, MC has been proposed to be a necessary and sufficient criterion for concept individuation and assembly at indefinite levels of perceptual abstraction, in a capacity termed *perceptual MC*. Perceptual MC is argued to be an innate and domain-specific competence in a dedicated music faculty (Rawbone, 2021). The status of MC within the theory of perceptual MC is distinct from that of generative theory, where MC is a non-necessary and non-sufficient feature in the construction of musical representations. In perceptual MC, two domains are thought to be privileged in the music faculty, harmony $\boldsymbol{H}$ and grouping $\boldsymbol{G}$, which are grounded by pitch and time regularity percepts, respectively. All other musical parameters emerge from these pitch and time regularity percepts (Rawbone, 2021). The co-occurrence of similarly stable elements within a single parameter (/domain) is termed uniparametric congruence (UC). UC is the state of affairs where stable elements in a single domain coincide with other stable elements of that domain, or conversely, where non-stable elements coincide with other non-stable elements of a domain. Therefore, to extend the definition of MC above, MC is the co-occurrence of similarly stable UC elements across different parameters (Rawbone & Jan, 2020). It should be noted that since the elements of the two domains, $\boldsymbol{H}$ and $\boldsymbol{G}$, are disparate phenomena, MC does not concern connections between elements (of $\boldsymbol{H}$ and $\boldsymbol{G}$) that are *actually* similar, but only *analogically* similar. MC connects UC elements (/features) that are similar only in terms of their degree of stability (i.e., their degree of UC), not their intrinsic structure.

The framing and treatment of MC relations in generative theories and generative computational implementations can be challenged on a number of grounds. It is questionable whether many of the well-formedness rules of GTTM-TPS, which are crisp, idealised and universal music perceptual or cognitive categories, can form the basis of graded and flexible mental representations (Benjamin, 1984; Peel and Slawson, 1984; Gjerdingen, 2007; Muns, 2014, 2015). Accordingly, it is doubtful whether such can be a foundation for MC interaction (Rawbone & Jan, 2020). For instance, metre and chords rarely exist as well-formed categories, being more commonly graded and non-isomorphic entities in many styles, historical periods and individual artefacts (Narmour, 1977; Gjerdingen, 1988, 2007; Rawbone & Jan, 2020). A tonic chord with an added seventh tone above the bass tone, such as a chord I^7 in C Ionian, i.e., {C, E, G, B}, may broadly function as a tonic chord set, i.e., {C, E, G}, yet it is not a pure (/well formed) tonic chord set, owing to the added B tone. The added tone changes the essence of the harmonic set as a whole and in turn changes the local or global interpretation of consonance or stability. In TPS, the problem of added tones or non-chord tones, termed “"distinctive pitch classes”, is dealt with as a factor of the “chord distance rule”' (Lerdahl, 2001, p. 55), which results in a faux graded categorisation of chord concepts. Indeed, since these distinctive classes are treated as separable and independent terms in the calculation, this results in fragmentation of the harmonic representation. These nuances must rather be represented within the harmonic entity itself, because they are actually intrinsic, not extrinsic, to the concepts as such. GTTM also involves faux graded harmonic (and metrical) representation by using tree-like elaborations in parent–child hierarchies to integrate harmonic entities. However, the combinatorial “expanding out” of symbols in this generative process arguably does not enable true gradedness, because it does not allow elaborations to yoke to underlying musical essences. As noted, it is the essence of the harmonic concept as such that requires faithful representation. More broadly, it is tendentious whether well-formedness rules are viable universal perceptual categories because they involve fairly high-level structures that can be decomposed into constituent elements, and it is arguably those constituent parts that are more likely candidates for perceptual universals. The lack of foundation for well-formed structures is an emphatic example of the symbol grounding problem, which concerns the difficulty of models in cognitive science and artificial intelligence to justify a symbol–world connection (Harnad, 1990). The validity of well-formed categories in GTTM-TPS and their basis as, or their connection with, innate archetypes is suggested or asserted, rather than explained and justified.

The other main rule-type of GTTM-TPS, the semiformal preference rules (/constraints), likewise poses a challenge for interpretation and implementation. The four central components that comprise the well-formedness rules and the interacting preference rules are organised in a circular, mutually-dependent architecture, where each rule presupposes the existence of others. Using this framework, it is an open question regarding how the system should be organised and parameterised for algorithmic implementation (Hamanaka et al., 2005, 2006, 2007; Hamanaka & Tojo, 2009; Marsden, 2010; Muns, 2014, 2015; Tojo, Marsden, & Hirata, 2018; Marsden et al., 2018). A further systemic limitation is that preference rules can result in incoherent representations because the rules are actually abstractions of more concrete underlying relational principles. The relatively fixed and brittle preference rules often conflict with each other because they do not coherently represent underlying MC forces. For example, in the GTTM-TPS program, grouping preference rules are formulated to cue the fixation of a single monistic well-formed metrical hierarchy. However, in contrapuntal textures, broadly construed, where several grouping streams occur simultaneously and with relative independence, the grouping preference rules should in principle cue multiple metrical forms. Yet polymetre is not accounted for in the rule system, presumably because it would create ill-formed hierarchies (i.e., resulting in paradox). Generative computational implementations that involve systems that automatically classify textures under a single monistic metre, such as Temperley (2001), Hamanaka et al. (2005, 2006, 2007), Hamanaka & Tojo (2009), Marsden (2010), Tojo, Marsden, & Hirata (2018) and Marsden et al. (2018), therefore avoid paradox but implicitly court systemic incoherence. Systemic incoherence between other preference rules, such as between grouping and metre or between harmony and grouping, is endemic in many of the generative models cited, which connect or disconnect features on the basis of abstract constraints, resulting in either over-generalisation or over-specialisation.

To address these issues of representation there is a need to look beyond finite well-formedness rules and relatively brittle preference rules, to capture the underlying graded MC relations that connect UC features. In the theory of perceptual MC, concrete metre is generalised to abstract metrical structure, allowing freer, non-constituent hierarchies at high levels of grouping and metrical structure (Rawbone, 2021). In this approach, texture is suggested to be regular, or well-formed, only at a low level, at beat or tactus levels, meaning only low-level units are fixed as basic incorrigible percepts. Ill-formed or pluralistic grouping and metrical structure may form at high levels, but always based on the low-level generators. Indeed, perception is free to generate ill-formed high-level structural representations but using the "building blocks" of fixed concrete low-level percepts. Similar depictions of high-level metrical freedom and plurality have been theorised in Cone (1968), Benjamin (1984), Povel & Essens (1985), Lester (1986) and Rosenthal (1992), but current symbolic and inductive models of harmony and grouping generally do not explore such structuring. The idea of a multiplicity of graded and unbounded relations is antithetical to received formalistic notions of metrical grids, feature templates, and statistically-learned patterns (e.g., Longuet-Higgins & Steedman, 1971; Leman, 1995; Temperley, 2001; Huron, 2006; Dhariwal et al., 2020). In formalising a coherent theory of grouping relations, then, it may be foundational that there are fixed low-level time-regular units, and that flexible formulation of UC groups is permitted at high levels of the ***G*** domain. *Mutatis mutandis*, there must be a distinction between low- and high-level concepts in the ***H*** domain, where pitch percepts are fixed at low levels, but high levels of harmony and tonality are free to generate graded, generalised and pluralistic UC forms (Rawbone, 2021.

The time-span reduction module of GTTM-TPS might be termed a meta-conceptual component, because it involves the framing and curation of events based on the well-formedness and preference rules of the other components. To continue the argument above, a difficulty with verifying and implementing time-span reduction is interpreting the circular rule system (Temperley, 2001; Marsden, 2010; Muns, 2015). Most computational implementations require some degree of human judgement in fixing the parametric rules to qualify salience in the time-span hierarchy (e.g., Hamanaka et al., 2006; Marsden et al., 2018). A response to the problem has been to develop a partially automated system of determining time-span significance by leveraging the influence of relative branch height in event tree hierarchies (Tojo, Marsden, & Hirata, 2018; Marsden et al., 2018). However, it may still be countered that time-span reduction is questionable in principle, because it nonetheless arbitrarily and incoherently construes MC relations based on a circular rule system design. By deriving time-span reduction from well-formedness and preference rules, it is not possible to determine MC relations without arbitrariness and systemic incoherence. More broadly, the lack of parsimony of the time-span reduction component diminishes the case that this module represents an actual property or process of perception or cognition. The implementation of time-span reduction involves considerable computational complexity, and so would be highly expensive in perceptual and cognitive resources (Marsden, 2010). Thus, it seems not to be a viable process for determining MC connections and individuating and assembling concepts. Toward constructing a system that accounts for MC interaction and avoids incoherence, inefficiency and arbitrariness, it is necessary to invoke relations between features directly, without meta-conceptual frames.

Prolongational reduction, which involves the hierarchical organisation of events based on harmonic stability (Schenker, 1935), is likewise a tendentious meta-conceptual component of the generative program (Gjerdingen, 2007; Marsden, 2010), and questionable as a basis by which to frame MC feature interaction. There are a number of implementational difficulties with this component which, following the line of critique played out above, largely reduce to the circular rule system design. In generative theory, prolongational reduction is framed as *the* output factor of musical perceptual experience. This focus, which places harmony, or harmonic hierarchies, as the essence of music processing, implicitly delegates other parameters to a less significant role in the experiential arena. A focus on harmonic prolongation leads to systemic incoherence because prolongation is just one parameter in the interacting constellations of parameters that inform concept construction. This bias is characteristic of both symbolic and inductive music theories and computational models, which often overplay the role of harmony or harmonic prolongation at the expense of other parameters, and frequently model harmony in isolation (e.g., Longuet-Higgins & Steedman, 1971; Leman, 1995; Chew, 2001; Bharucha, 2009; Rohrmeier, 2011). Problematically, harmony is incorporated as both an input factor (e.g., "harmonic length" preference rule (Lerdahl & Jackendoff, 1983, p. 84)) and an output factor in GTTM-TPS, combining innate and acquired harmonic constraints, and conflating bottom-up and top-down processes, and so making it difficult to position harmony, and harmonic concepts in general, within the sequential steps of any proposed algorithm.

Complex harmony-related concepts, such as harmonic change, harmonic rhythm, etc., either involve *ad-hoc* or *post-hoc* preference rules (/constraints), including the “harmonic length” preference rule of GTTM, or are not incorporated into the theoretical or implementational system. These concepts often do not connect coherently with the general harmonic prolongational form, resulting in systemic fragmentation, and are thus a challenge for computational implementation. It is noteworthy that while in some other areas of music research complex harmonic concepts are considered decisive factors for representation and interpretation (e.g., Lester, 1986; Mirka, 2009), in generative approaches they have minimal or no significance. This position can be explained by a lack of a formal framework to incorporate complex harmonic concepts, possibly because a formal representation of harmony is itself at present largely inscrutable. While harmonic prolongation is an intuitively reasonable, although limited, meta-conceptual metaphor, it may ultimately be a misleading abstraction in the modelling of harmony and tonal structure (Narmour, 1977). To address the issue of formalising harmony and tonal structure – and *mutatis mutandis*, formalising grouping and metrical structure – there must be a more direct and coherent measure of the connection between musical terms. In general, (fuzzy-)logical techniques have been given only limited attention in music research and the cognitive sciences for the explanation of musical structure (Cádiz, 2020). The present model aims to show that (fuzzy-)logical and set-theoretical connections can determine the boundedness and interaction between musical terms through time, and thus may be used to fix concepts in perception. Accordingly, the present framework may suggest a revision of the traditional notion of harmonic prolongation.

The theory of harmonic prolongation in TPS leverages the idea of pitch space, which involves posited mental spaces for mapping pitch, chord and key relationships. A focal point of pitch space theory in TPS is the chord distance rule (discussed above), which calculates mental “distances” between pitches, harmonies and keys, usually in the context of a central tonic (pitch, chord, or key). Mental distance calculations are based on the principle of the shortest path (Lerdahl, 2001, p. 55), which means calculating distances according to the most direct route. While an ostensibly useful system for quantifying pitch and harmony relatedness, pitch space theory has little empirical support (Lerdahl, 2001). TPS itself involves highly complex mappings of inter-dimensional connections that would require extensive perceptual and cognitive resources. To enable (faux) graded harmonic connections between (crisp) well-formedness rules, additional well-formedness rules and *post-hoc* preference rules are incorporated. For instance, along with a formalisation of well-formed mental spaces in TPS (e.g., diatonic space, hexatonic space, etc.), surface dissonance constraints are appended to qualify pitch-space distance calculations. The use of *post-hoc* rule structures is one of a series of corrective strategies in generative theories that belie systemic incoherence. In general, representational music theories, including spatial and geometric models, are often highly complex and contradictory (e.g., Riemann, 1905; Lerdahl, 2001; Tymoczko, 2012). The validity of spatial models such as pitch space theory (e.g., Lerdahl, 2001), two-dimensional graphs such as the *Tonnetz* (e.g., Riemann, 1905)), topographic models (e.g., Mazzola et al., 2002), and voice-leading geometries (e.g., Tymoczko, 2012), can be disputed based on computational complexity (Marsden, 2010), inefficiency, structural inelegance, paucity of consilience between individual models, and poverty of empirical support (Lerdahl, 2013). The present work appeals to a seemingly more elegant basis for harmonic categorisation, *harmonic root support*. Harmonic root support involves classifying pitch sets into chords by virtue of their combined support for a common root tone (Terhardt, 1982; Parncutt, 1988; Milne, 2013). The theory of harmonic root support is analogous to the theory of pitch perception, since finding the root of a chord involves a similar process to determining the fundamental frequency or virtual frequency of a complex tone. An advantage of root support as a basis for harmonic categorisation is that it requires only a single dimension (a single harmonic function) to unify a harmonic set, and so is easily applied to many other harmony-related concepts. It is by virtue of this principle that the graded structure of harmony will be shown to be intrinsically coherent. A generalisation of this notion (*generalised harmonic root support*, Section 3) also enables externally coherent connections between harmonic sets of domain $\boldsymbol{H}$ and broader coherence with domain $\boldsymbol{G}$.

In summation, the lack of parsimony of generative theories, and the systemic circularity and fragmentation in and between the four main components (grouping, meter, time-span reduction and prolongational reduction), suggest they are unlikely to have a basis in perception or cognition. It is questionable whether the generative research program in its current form reveals the nature of MC interaction or the process of concept individuation and assembly.

1.2 Fuzzy Music Generation

In this subsection, fuzzy set theory and fuzzy logic literature are reviewed as tools for determining sets (/concepts) and relations in and between the musical domains. Fuzzy sets are an extension of classical (/crisp) sets in that they generalise the truth value of a set using graded membership functions. Similarly, fuzzy logic generalises the inference connectives of classical logic (i.e., OR, AND, NOT, etc.) by including graded connectives. That is, fuzzy sets and connectives involve graded truth values and truth functions with membership values that lie between 0 and 1 (Zadeh, 1965). Fuzzy techniques have been used in recent years in a diverse pool of disciplines and technologies to model physical processes and human decision-making, particularly in machine learning (e.g., Alsinet & Godo, 2000; Fitzgerald et al., 2004; Armengol et al., 2015; Hüllermeier, 2015), fuzzy control systems engineering (e.g., Zhang & Liu, 2006; Azad & Shukla, 2021), various sub-disciplines of linguistics (e.g., Cock et al., 2000; Sun et al., 2002; Novák & Perfilieva, 2004; Gupta et al., 2018), visual perception (e.g., Lakoff, 1987; Paz et al., 2019), and other areas of cognitive science (e.g., Massaro & Cohen, 1993, Yahia et al., 2012).

Important for modelling the interaction between musical categories is the idea of a fuzzy relation, which concerns the graded interconnectedness between sets. While classical relations between sets are crisp and bivalent, the notion of a relation is generalised to involve graded (fuzzy) interconnections (Zadeh, 1965). Classical sets and relations, and fuzzy sets and relations are suited to different applications. Some phenomena in the world can be best represented using classical sets and relations, while other phenomena are more optimally captured using fuzzy sets and relations. For instance, human familial biological relations are most suitably defined through classical relations. Biological sibling–sibling relations or parent–child relations involve crisp bivalent functions, where membership is *either* 0 or 1. By contrast, human friendliness sets and friendship relations involve fuzzy sets and fuzzy relations, which lie *between* 0 or 1, because there are varying degrees of friendliness and friendship (Bach, 1964). There are several types of relations used in fuzzy set theory and fuzzy logic, such as resemblance, partial order, equivalence, etc., which have various properties, such as reflexivity, symmetry, transitivity, etc., and these are often depicted in matrix form. The properties of fuzzy relations are conventionally defined quite rigidly, although recent work in fuzzy set theory has weakened some properties to enable increased gradedness and generalisation. For instance, the properties weak fuzzy reflexivity (w-reflexivity) and weak fuzzy symmetry (w-symmetry) are more generalised forms of fuzzy reflexivity and fuzzy symmetry, respectively, giving them wider scope (Yeh, 1973; Gupta & Gupta, 1996; Chon, 2017). In fuzzy logic, implication is often modelled as the boundedness or overlap between fuzzy sets, having a similar function as the intersection connective, involving any t-norm operator (Zadeh, 1965). Fuzzy relations and their properties form the core of FRMP, presented in Section 5.

It may be thought that there would be widespread acceptance of the common-sense notion that music concepts (/sets) and their relations are fuzzy. However, in music theory, crisp sets and their relations are more commonly adopted to depict musical entities and their interconnections (e.g., Forte, 1973; Buchler, 2001; Kuusi, 2001; Mazzola et al., 2001; Tymoczko, 2022). There is only a modest quantity of models that utilise fuzzy sets and fuzzy logic in music research, and these include systems addressing signal processing issues (e.g., Malcangi, 2008; Gonzalez-Inostroza et al., 2015), fuzzy music emotion classification (e.g., Yang et al., 2006; Suitar, 2010; Kumar et al., 2015; Lucas et al., 2017; Hasanzadeh et al., 2019; Kasinathan et al., 2019), and fuzzy memory and music information retrieval (e.g., Weyde & Dalinghaus, 2001; Monti & Sandler, 2002; Deliège & Pedersen, 2007; He & He, 2019). The quantity of studies that use fuzzy techniques to address issues of music perception and generation, the concerns of the present article, is limited (Cádiz, 2020).

An early fuzzy music generation model, Elsea (1995), involves a system for representing fuzzy pitch sets in connection with other basic parameters. It introduces fuzzy sets and fuzzy logical operators for chord inversions, added chord tones and fuzzy logical chord progressions. However, the content of chord sets and the connectives of progressions emerge from the high-level control system, which computes traditional rules based on logic rooted in abstract musical–linguistic commands, rather than explicating the structure of bottom-up generative principles. As such, it does not explain basic music concepts as sets and partition structures that determine the innate, universal and bottom-up categories used in perception. Tokumaru et al. (1998) introduce systems of automatic harmonisation of melodies using fuzzy relational operators to contextualise parameters, involving a genetic algorithm that learns optimal harmonisation solutions. Similar to the criticism of Elsea (1995), a limitation of these models is that the central basic and complex concepts that are vital for music perception, relating to harmony and grouping information, are assumed and asserted as part of the high-level rule systems, rather than given formal explication and representation in terms of low-level fuzzy elements, sets, partition structures, connectives and relations. However, Tokumaru et al. (1998) incorporate connections between parameters that are based on, or analogous with, some preference rules of GTTM, highlighting such factors as tonic salience, note length, metrical depth, chord function and chord progression, in the determination of representations.

Yilmaz & Telatar (2010a, 2010b) examine the cognition of two-part counterpoint and four-part harmonisation rules as practised in the eighteenth century. They introduce definitional fuzzy equations for pitch motion, interval size, consonance and dissonance, harmonic sets and voice-leading rules. The fuzzy voice-leading rules pair pitch events in time undergoing serial motion in counterpoint, providing a fuzzy formalisation of parallel, similar, oblique and contrary motion. Similar to the criticism of the above fuzzy models, Yilmaz & Telatar (2010a, 2010b) focus on high-level rules presumably used in cognition, not fundamental processes and properties central for representation of MC relations in perception, and which are arguably intrinsic to concept formation. For instance, the models do not show how harmony is perceived by virtue of the interaction between interval content and inversion (cf. Parncutt 1988), among other intra-parametric interactions. As such, fuzzy chord sets are not defined in terms of their interval content, inversions and partition structures, but emerge simply as a by-product of the high-level contrapuntal rules. Accordingly, the systems in Yilmaz and Telatar (2010a; 2010b) do not deal with the perception of harmony in general. It should be uncontroversial that we at least *partly* perceive a chord by virtue of such qualities as its interval content and inversion (Parncutt, 1988). A focus on high-level traditional contrapuntal rules, while important for modelling the specific style in question, is of limited use for understanding universal processes of perceptual concept individuation and assembly, which seem to be the primary activities carried out in the music perception module, or music faculty. In general, many of the models discussed do not provide a basis by which we can explore MC interaction for the representation of universal basic and complex concepts and their interconnections.

While lacking a satisfactory framework, it can be argued that fuzzy techniques may still in principle be used to simulate universal relations between musical features involved in the fixation of concepts. Fuzzy techniques provide tools for addressing the symbol grounding problem (defined above). They are able to model concept individuation and assembly more faithfully than generative and associative–statistical systems, such as deep learning algorithms (cf. Civit et al., 2022; Golik et al., 2012; Dhariwal et al., 2020), because they can model the interaction of graded musical entities directly, avoiding meta-conceptual frameworks, and do not rely on the various problematical desiderata of concept-learning, such as dataset assimilation. Indeed, the notion of concept-learning may be questionable in principle because inductively-generated concepts still require some form of prior representation, or pre-interpretation, to enable recognition or confirmation of concepts during the process of "induction" (Fodor, 2008). For instance, the architecture of a connectionist system (supervised or unsupervised) must involve various forms of parameterisation or "framing" at the outset that necessarily inculcates myriad types of privileged information that substitute for innate human knowledge – which is to say that the notion of concept-learning is a regress argument (Fodor, 1975, 2008). Thus, there is a compelling analytical argument for a model of perceptual processing that is intrinsic, innate and universal, not based on obscure individuation mechanisms and empirically-guided concept-learning.

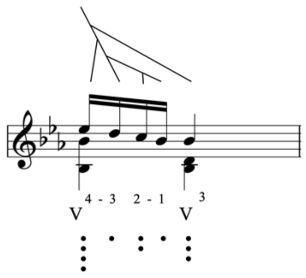
1.3 Rationale for Fuzzy Relational Music Perception

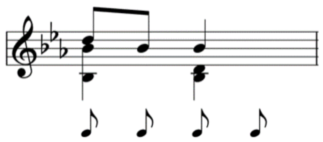
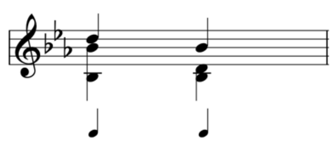
This subsection develops the case for an innate graded relational system for fixing music concepts, setting out a principled basis for FRMP. A dedicated music-perceptual module, a music faculty, seems to be requisite for coordinating MC interconnections. The existence of some form of music faculty has support across a number of disciplines, such as neuroimaging (e.g., Griffiths & Frackowiak, 1998; Patterson et al., 2002; Zatorre et al., 2002; Tavalage et al., 2004; Yost, 2009; Oxenham, 2012; Nunes-Silva & Hasse, 2013; Norman-Haignere et al., 2015), neuropsychology (e.g., Peretz & Coltheart, 2003; Peretz & Zatorre, 2005; Peretz, 2006, 2009), and philosophy (e.g., Fodor, 1983; Schneider 2011). In music theory and cognitive science, Patel (2008), Katz and Pesetsky (2011) and Lerdahl (2013) invoke weak forms of music faculty, which comprise shared capacities or shared neural structure between music and natural language systems, and possibly other perceptual or cognitive systems. However, concept assembly in music using MC involves dissimilar entities and processes from those of natural language, so the extent of shared neural circuitry for these activities may be questioned *prima facie*, even if some operations are broadly parallel (Rawbone, 2021).

Underpinning concept assembly is the principle of compositionality, which is the idea that basic atomic percepts and low-level concepts are combined to form hierarchies of complex concepts that preserve the lower-level rules or semantic structure of the basic percepts and low-level concepts (Fodor, 1975, 2008; Partee, 2011). Indeed, a central function of the music faculty seems to be to construct complex conceptual hierarchies from basic atomic percepts in a way that is compositional. It can be presumed that the fixation of basic percepts is grounded by sound-wave-to-neural transduction during initial processing in the cochlea, and that these percepts are subsequently presented to the auditory cortex, the likely location of the music faculty (Cohen et al., 1995; Justus & Bharucha, 2002; Bharucha, 2009; Rajendran et al., 2018; Rajendran et al., 2000; Rawbone, 2021). In Rawbone (2021), it is theorised that *pitch* and *low-level time-regular units* are the basic percepts. A basic pitch percept is individuated from a harmonically complex function that is automatically encapsulated by reference to support of a fundamental frequency or integer multiples of a fundamental (virtual pitch) (Deutsch, 1998; Justus & Bharucha, 2002; Bharucha, 2009; Peretz, 2006, 2009). A basic, low-level time-regular percept is coined through innate entrainment to the beat or tactus level through time (Rajendran et al., 2000; London, 2004). Since the basic percepts are thought to be universal, learning probably has a limited role in their coinage. Any influence of learning from the environment in this process must occur early in life and developed to proficiency in the first few months (cf. Justus & Bharucha, 2002; Bharucha, 2009), since newborn and infant perception and comprehension of music concepts, and their relational interaction, is highly developed (Peretz, 2006, 2009).

The basic percepts are privileged in the music faculty, since all other parametric features (/complex concepts) emerge from them. Complex concepts are formed from the basic percepts of the two fundamental domains, harmony $\boldsymbol{H}$ and grouping $\boldsymbol{G}$. As noted, the principle of compositionality is axiomatic for concept construction. Accordingly, the basic percepts are not revised through top-down processing, which would result in incomprehensible and incoherent complex concepts (Fodor, 2008). Basic percepts are incorrigible (non-revisable), but complex concepts are continually revised in a chain of causation that ultimately stems from composers and cultures (Jan, 2007). Complex concepts, such as harmony, grouping, tonal structure and metrical structure, are continually re-represented during the unfolding of events, since emerging concepts require adaptive revision to cohere to the unbounded interaction at various levels of abstraction. A framework is required to faithfully describe unbounded MC interaction. Pitch and time-regular units must be irreducible percepts grounded by transduction, but flexibly connected and rearranged into complex concepts for indeterminate and graded high-level conceptual hierarchies. In $\boldsymbol{H}$, tonality is a higher-level generalisation of harmony (*inter alia*), and harmony is a generalisation of pitch percepts (*inter alia*). Likewise, in $\boldsymbol{G}$, metrical structure is a higher-level generalisation of grouping (*inter alia*), which, in turn is a higher-level generalisation of low-level time-regular units (*inter alia*). The qualification "*inter alia*" (meaning *among other parametric features* in these contexts) is important because complex concepts must involve information based on the interaction between the domains. Indeed, low-level complex concepts (e.g., harmony and grouping), mid-level concepts (e.g., harmonic rhythm) and high-level concepts (e.g., tonality and metrical structure) are formed from relations between $\boldsymbol{H}$ and $\boldsymbol{G}$.

A central thesis of this article is that for concept construction perception connects $\boldsymbol{H}$ and $\boldsymbol{G}$ based on MC relational strength. Perception uses innate knowledge of MC relations to determine similarly stable UC features between the domains. Indeed, since features are variably abstract, hierarchical and graded, and emerge in an indeterminate process, perception must have innate knowledge of graded MC interconnections to determine similarly stable features to individuate and assemble concepts. In the following discussion, Ex. 1 (a)–(c), bar 2 of Mozart's Piano Sonata No. 4, i (K. 282), will be used to demonstrate innate knowledge of MC relations. Ex. 1 (a) involves the original manuscript notation, with added Roman numeral harmony analysis and metrical dot structure. Ex. 1 (b) is the abstract reduction at the quaver level, and Ex. 1 (c) is an abstract reduction at the crotchet level.



(a)

(b)

(c)

Example 1: (a) Mozart Piano Sonata (K. 282 i, b. 2); (b) reduction at quaver level; (c) reduction at crotchet level (based on Lerdahl, 2001, pp. 152–153).

Referring to Ex. 1 (a)–(c), Lerdahl (2001, pp. 152–153) builds a case that a hierarchical analysis (involving various levels of abstraction) is critical for perception to determine unstable tones from stable tones. Perception fixes the harmonic background as an abstract B♭ chord so that the E♭ tone in the melody can be categorised as more unstable than the higher-level D tone, and the C tone can be classified as more unstable than the following higher-level B♭ tone. That is, it is requisite that the higher-level B♭ harmony is an abstract category for the D and B♭ tones to be perceived on higher hierarchical levels than the E♭ and C tones. This analytical principle generalises to all musics, and so seems to be an innate property of music perception. (See Deutsch (2013b) for a review of psychological models of hierarchical analysis.)

Lerdahl's argument can be extended to cover MC relations. A central premise of FRMP is that perception requires implicit knowledge of graded MC relations between ***H*** and ***G*** domains to categorise the hierarchically structured, graded and variable concepts that emerge in real-time. To abstract a B♭ major harmonic category in Ex. 1 (a)–(c), perception must have implicit knowledge of MC relations to make sense of the unfolding relational connections between domains. The stable UC harmonic category of B♭ major must be relationally connected in perception to stable UC grouping (and metrical levels, more abstractly) at the quaver level (Ex. 1 (b)) and crotchet level (Ex. 1 (c)), because it is by virtue of the MC relations that the concepts are individuated as such. That is, the stable UC harmonic category is only classed as B♭ because it forms a graded MC relation with stable UC levels of grouping structure. Such informational interconnections emerge arbitrarily in real-time, and so implicit knowledge of MC relations must be used to interpret them and build concepts accordingly. As a *reductio* argument, if real-time perceptual interpretation using implicit knowledge of MC relations was not used, the harmony on beat 1 in Ex. 1 (a)–(c) could be categorised as various other possible chords, such as an E♭ chord in second inversion (with a missing third), which corresponds to the sensory data presented at this time-slice. Only with innate knowledge of MC relations, incorporated in real-time, can this time-slice be conceptualised under the general B♭ major harmonic category. Innate sensitivity to the emerging MC relational context is thus essential for the appropriate intentionally-constructed conceptual categories to be assigned. The upshot here is that since feature interaction is a fluid process, where interconnections emerge differentially between features on every occasion – where graded emergent information in ***H*** (and graded tonal information, more abstractly) is continually and arbitrarily cross-referenced with graded emergent information in ***G*** (and graded metrical structure, more abstractly) – intrinsic and innate knowledge of MC relations must be requisite in perception. The present approach assumes a distinctly different system to the rigid and predetermined preference rules of generative theories, where well-formed representations are preferentially connected to each other at the outset. From the present perspective, it is not viable, as in generative theories, to use a series of predetermined preference rules to classify interaction between features for concept construction, since any such rules would be brittle and incoherent with the indeterminate fuzzy relational possibilities that emerge in real-time.

It follows that any sets based on concrete pitch or time-regular units, such as scales, modes, or rhythmic groups, cannot in isolation individuate complex concepts such as harmony, grouping, tonality, or metrical structures, among others. While such concrete sets as scales and modes (e.g., Ionian, Dorian, Phrygian, Lydian, Mixolydian, Aeolian, Locrian) are associated with tonalities, through perception, cognition and socio-cultural conditioning, they are not causally sufficient for individuating complex concepts in perception, since they involve only a single parameter, i.e., intervallic content or pitch chroma. Rather, in Ex. 1a–c, the key chord tones of the B♭ major scale (tones B♭, D and F) emerge in real-time by virtue of the relational strength (MC interaction) between $\boldsymbol{H}$ and $\boldsymbol{G}$ at various structural levels. Also, non-chord degrees of this scale (tones C, E♭, G, and A) form owing to a differentially lower MC relational strength between the domains (Section 5). It seems that notions such as scale are ephemeral and indefinite because they emerge as by-products of MC relational strength during concept-building. While the socio-cultural commonality of particular serial sets, including scales, memes and schemata, are partly a product of the top-down causation of culture and cognition (Jan, 2007), such sets may have limited significance in the perceptual realm. Thus, the overarching moral to be drawn from Ex. 1a–c is that the parameters involved and concepts created in music are generated exclusively by interaction between the two fundamental domains, $\boldsymbol{H}$ and $\boldsymbol{G}$. Also, the individuation of concepts in perception has been argued to be necessarily based on innate knowledge of MC. In the following subsection, fuzzy *resemblance* and *implication* relations will be shown to be the necessary and sufficient formal mathematical relations in and between the domains to individuate fundamental and universal music concepts.

1.4 Theoretical Components of Fuzzy Relational Music Perception

The above discussion suggests that abstract, hierarchical and graded MC relations in and between the domains are represented during perceptual processing. This subsection sets out the main theoretical components of FRMP, outlining the principles used to construct both domains and the fuzzy relational mathematics used to model domain interaction.

Domains $\boldsymbol{H}$ and $\boldsymbol{G}$ concern disparate informational phenomena. Domain $\boldsymbol{H}$ is determined by *generalised harmonic root support* with respect to a main tonic pitch of a tonic chord. In constructing a domain, the root of a central chord generalises over all harmonic sets of $\boldsymbol{H}$ for a particular key or tonal area, quantifying UC in that domain. Generalised harmonic root support in $\boldsymbol{H}$ is mirrored, *mutatis mutandis*, in $\boldsymbol{G}$ by *generalised time regularity*, which governs over all grouping sets in $\boldsymbol{G}$, and likewise quantifies UC in that domain. Two types of fuzzy relations are used in FRMP: fuzzy resemblance (also termed tolerance or compatibility (Ross, 2010)) and fuzzy implication. Fuzzy resemblance depicts the graded and variably abstract interconnections between the domains and fuzzy implication determines the boundedness (/overlap) between elements through time in one or both domains. Representations of fuzzy resemblance and implication relations enable coherent and elegant formulations of concepts such as harmonic progression, harmonic change, harmonic rhythm, tonality, metrical structure, grouping, passing tones, etc. The two fuzzy relations are formally defined as follows:

1. **Fuzzy resemblance relation**. Features of $\boldsymbol{H}$ form a fuzzy resemblance relation R_R with features of $\boldsymbol{G}$ through time in a $\boldsymbol{H} \times \boldsymbol{G}$ Cartesian product space, modelling MC interconnectivity. R_R comprise the properties of weak reflexivity, weak symmetry and antitransitivity.

2. **Fuzzy implication relation**. Elements of domains form a fuzzy implication relation through time in a $\boldsymbol{H} \times \boldsymbol{G}$ product space. Fuzzy implication relations concern the fuzzy overlap or boundedness between set elements in the product space in one domain (R_M) or both domains (R_I) through time, involving a UC or MC connection, respectively, and leveraging the intersectional min operator.

These fuzzy relations enable the coherent integration of the two disparate domains. As noted in Section 1.3, since the domains are informationally distinct, graded, abstract and variable through time, without innate knowledge of these relations it would be impossible to glean informational content from the whole unbounded and combinatorial musical landscape, and so it would be impossible to individuate and assemble concepts. In this sense, the music faculty, which is the entity that FRMP simulates, and the incoming musical information, are analogues of public-private encryption algorithms and public data. Music faculties are like systems with private keys that allow only us, as humans, to decode the public emergent fuzzy relations between the disparate domains of music presented to us by composers and cultures. That is, the decoding of the relational interconnections between the domains using fuzzy relations allows the intentionally formulated concepts of composers (and cultures) to be individuated and assembled, amounting to an interpretable musical thought language (Rawbone, 2021). (Note that the composer of strings in the musical language and the perceptual decoder of that language can be, and in all cases of composers necessarily are, one and the same individual.) Fuzzy set-theoretical resemblance (R_R) will be shown to decode the analogical connection between domains, and fuzzy-logical implication will be used to decipher the continuation of concepts in a single-domain, i.e., $\boldsymbol{H}$ or $\boldsymbol{G}$ (R_M), or in both domains combined, i.e., $\boldsymbol{H}$ and $\boldsymbol{G}$ (R_I).

While these relations form the main components of FRMP, they rest on a significant body of definitional theorisation and formative equations that are set out prior to presentation of the core relations. To provide support for FRMP, each stage in the explication of relations and relational hierarchies is illustrated with musical examples which reveal the action of perception in representing musical structure. Ancillary theory is appended to account for pluralistic hierarchies, involving two or more independent but internally coherent conceptual hierarchies that are superimposed onto a single musical surface, and where any relationship between the hierarchies is to a large extent opaque. While contrapuntal musical genres are paradigm cases of pluralistic relational hierarchies, many musics of the world incorporate such structuring, and so the capacity to generate them seems to be innately endowed.

1.5 Article Structure

The remainder of the article is structured as follows. Section 2 illustrates that fuzzy sets are partitioned along $\boldsymbol{H}$ and $\boldsymbol{G}$ domain continua using intersecting triangular membership functions (triangular norms, or t-norms). The main operators incorporated for building fuzzy sets, domains and implication relations are outlined. Section 3 shows how harmony is categorised using the theory of root support, and that the generalisation of this principle determines the partition structure of $\boldsymbol{H}$. Section 4 shows that time regularity is the basis of grouping and builds a partition structure in $\boldsymbol{G}$ through the generalisation of time regularity. Section 5 is the central theory of this article, positing that the interconnectedness in and between $\boldsymbol{H}$ and $\boldsymbol{G}$ through time is represented in perception as fuzzy resemblance and implication relations, which enable the individuation and assembly of concepts. Supplementary theory of plural relational hierarchies is incorporated to characterise polyphonic structure, broadly conceived. Section 6 evaluates the overall significance of FRMP for concept construction and contextualises the model in terms of inductive and symbolic frameworks. The limitations of the model are its informal components, such as the manual determination of membership functions from partition structures and constraints, and the peripheral treatment of chromaticism and modulation. Future developments may include examining the role of plural hierarchies as a factor of cognitive complexity and providing computational implementations of the full model.

2 Fuzzy Musical Sets, Domains and Operators

This section lays out the fundamental mathematical principles of fuzzy musical domains, sets and operators. The ground for a fuzzy perceptual framework is prepared by providing an overview of classical set theory and classical logic. Crisp sets, connectives and operators are important foundations because they are special cases of (and subsets of) fuzzy sets, connectives and operators, and required for concretisation and consolidation of concepts. However, it is shown that crisp characteristic functions are limited in that they cannot show the common graded structure of sets in $\boldsymbol{H}$ and $\boldsymbol{G}$, or the graded interaction between domain sets. Fuzzy approaches have been used to model real-world problems and decision-making on account of their ability to capture gradedness in both external processes and conceptual representations (Ross, 2010), and are here posited to be more precise in representing music sets, connectives and operators. Fuzzy sets are thought to be constructed through time *on the occasion of sense data*, to echo Descartes' classic epithet. The elements of sets are the automatically transduced and coined basic percepts of pitch and low-level time regularity. There is an overview of the main operators used in fuzzy set theory and logic, i.e., min, max, $1 - \mu_X(u)$, and Mamdami implication, which enable the construction and interaction of sets in ***H*** and ***G***.

2.1 Classical Set Theory and Logic

In classical crisp set theory, a set, say X in the universe of scope $\boldsymbol{U}$, is defined in terms of its finite elements or members (i.e., x_n, where $X = \{x_1, x_2, \ldots, x_n\}$). The integer set notation {1,2,3,4,5,6,7} can be used to represent heptatonic scales mapped to the 12-tone chromatic scale, i.e., {C, C♯/D♭, D, D♯/E♭, E, F, F♯/G♭, G, G♯, A♯/B♭, B}, based on the theory of harmonic root support presented in Section 3. For instance, the tonality and modality of A Ionian is associated with the Ionian scale [A, B, C♯, D, E, F♯, G♯]. Thus, using integer set notation, a root A major chord set A in the Ionian scale is defined as follows:

$$A = \{1,3,5\}$$

In crisp set theory, a given element x either belongs to a set, say X, or does not, i.e., $x \in X$ or $x \notin X$. A Boolean characteristic function χ_X describes membership of X, where χ_X is either 0 or 1. Accordingly, given a chord set C, a musical pitch element x has the following Boolean characteristic function:

$$\chi_C(x) = \begin{cases} 1, & x \in C, \\ 0, & x \notin C. \end{cases}$$

Using the key and scale of A Ionian and integer set notation, integer 3 (pitch C♯), corresponding to the 3rd step of the scale set, is normally considered a member of the A major chord set A (*ceteris paribus*), and so has a characteristic function χ_A of 1. The issue with this representation is that such a depiction does not show possible degrees of membership of set elements. Also, while an ostensibly useful framework, crisp set theory is a limited picture of the harmonic landscape, since pitch content (pitch chroma) does not exhaust the content of graded and complex conceptual harmonic representation. In perception, harmony also has the properties of intervallic content, inversion, spacing and doubling (*inter alia*), and its graded elements vary through time (Parncutt, 1988). In this respect, crisp set membership and characteristic functions are limited approaches to modelling musical concepts, because they ultimately capture only a single parameter (pitch chroma).

The main operators that form the core of classical set theory and logic are intersection (/conjunction), union (/disjunction), complement (/negation) and implication. To take the first, intersection, concerning crisp sets in a universe of scope $\boldsymbol{U}$, the intersection of elements (u) of two abstract classical sets A and B is defined by the following expression:

$$A \cap B = \{u \in \boldsymbol{U} \mid u \in A \text{ and } u \in B\}$$

The union of two abstract classical sets A and B in the universe of scope $\boldsymbol{U}$ is defined by the following expression:

$$A \cup B = \{u \in \boldsymbol{U} \mid u \in A \text{ or } u \in B\}$$

Since in perception the interaction between domains ***H*** and ***G*** is graded, the aggregation of sets by crisp intersection and union do not allow capture of the possible connections that emerge. A further standard operator, material implication, is used in crisp logic, involving an implication relation R (→) between sets. For abstract sets A and B in the following, A entails B:

$$R = A \rightarrow B = \neg A \cup B$$

Prima facie, crisp material implication seems to have no place in music perception, since events in musical structure do not logically entail each other. Also, material implication would require that the semantic criteria of sets fall under well-defined premises, which does not seem to be the case in music. In the unfolding of music events, implication can, however, be captured by the graded boundedness or overlap between elements of sets through time, which requires a representation of the graded relations between terms. Crisp complement (/logical negation) likewise seems to be inadequate for modelling music perception because crisp sets, such as crisp pitch or chord sets, virtually never have absolute complements in music, but involve graded complements, or graded negations. Crisp complement is expressed in the following formula, which is read as u is *not* an element of set A in the universe of discourse $\boldsymbol{U}$. The A with a bar symbol above refers to complement, and broadly corresponds with the negation operator of classical logic, i.e., $\neg A$.

$$\overline{A} = \{u \mid u \in \boldsymbol{U} \text{ and } u \notin A\}$$

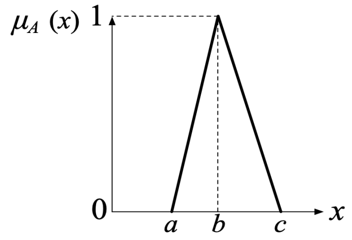
2.2 Fuzzy Triangular Sets and Domain Partition Structures

Fuzzy triangular sets and domain partition structures are fundamental to FRMP. Fuzziness, or gradedness, is sometimes problematically conflated with uncertainty and vagueness. To consider the first characterisation, uncertainty, it should be noted that probabilistic techniques and fuzzy techniques are mutually exclusive notions (Klir and Yuan, 1995). With respect to the second characterisation, vagueness, considering fuzzy processing as imprecise is antithetical to the innate and intrinsic capacity to perceive precisely graded musical concepts. It is therefore preferable to view fuzzy sets and relations as involving *precision gradedness*, to avoid confusion with probability or possibility theory, and to limit association with vagueness and imprecision.

A fuzzy membership function maps elements in a domain of discourse to a value between 0 and 1. Each element (u) of a universe of discourse $\boldsymbol{U}$ belongs by degree to some fuzzy set X, depicted with a graded membership function μ_X (u), expressed in Eq. (1).

$$\mu_{\underset{\sim}{X}}(u) \in [0,1], \text{where } \underset{\sim}{X} = \left\{(u, \mu_{\underset{\sim}{X}}(u)) \mid u \in \boldsymbol{U}\right\} \quad (1)$$

There are a number of standard membership functions used in fuzzy logic, such as triangular, trapezoidal, Gaussian, sigmoid, etc. (Klir and Yuan, 1995). Non-linear membership functions, particularly bell-shaped functions, may usefully characterise several musical phenomena, such as the added tones of chords, but are computationally expensive and may not effectively depict harmonic root support, so are not explored here. Triangular membership functions, i.e., triangular norms (t-norms) and triangular conorms (t-conorms) (also termed S norms), are used to generate the domain structure of $\boldsymbol{H}$ and $\boldsymbol{G}$, employing the min and max operators (Zadeh, 1965) on sets and relations (Section 2.3). T-norm and t-conorm membership functions are represented as linear, although domains $\boldsymbol{H}$ and $\boldsymbol{G}$ are actually based on non-linear logarithmic scales. Fig. 1 shows (a) graphical and (b) algebraic representations of a t-norm function for a fuzzy set A with membership function μ_A (x).



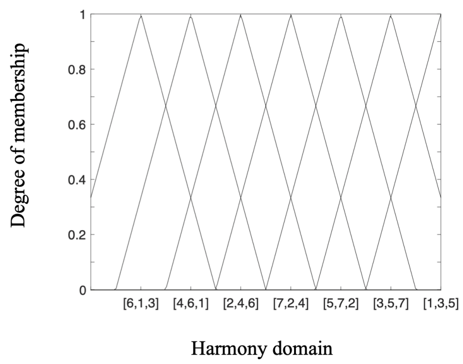
(a)

$$\mu_A(x) = \begin{cases} 0 & \text{if} \quad x \leq a \\ \dfrac{x-a}{b-a} & \text{if} \quad a \leq x \leq b \\ \dfrac{c-x}{c-b} & \text{if} \quad b \leq x \leq c \\ 0 & \text{if} \quad x \leq c \end{cases}$$

(b)

Figure 1: (a) Graphical and (b) algebraic representations of a fuzzy triangular membership function (Yahia et al., 2012).

Fuzzy sets have the usual properties of crisp sets: associativity, distributivity, idempotency, identity, transitivity and involution. However, the law of identity, law of excluded middle and law of contradiction do not hold for fuzzy sets (Zadeh, 1965). Both domains, ***H*** and ***G***, are proposed to be partitioned by t-norm sets that have two points of intersection, cutting the sets into three equal parts along the degree of membership axis (Fig. 2 and Fig. 3). Domain ***H*** will be shown in Section 3 to be based on generalised root support, constructed from hierarchical levels of perfect unisons, perfect fifths and thirds (major). Domain ***G*** will be shown in Section 4 to be based on generalised time regularity, constructed from fuzzy regular hierarchical durations. (The relative duration of grouping level sets in ***G*** are as follows: maxima = 8; longa = 4; breve = 2; semibreve = 1; minim = .5; crotchet = .25; quaver = .125; semiquaver = .0625.)



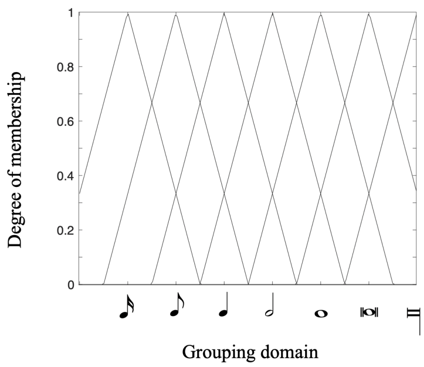


Figure 2: Triangular membership function partitioning of the *H* domain (set to [1,3,5]).

Figure 3: Triangular membership function partitioning of the *G* domain (set to longa).

2.3 Basic Fuzzy Operators

Basic fuzzy operators min, max and $1 - \mu_X(u)$ (Zadeh, 1965) are used in FRMP to determine intersection, union and complement, respectively, although other t-norms, t-conorms and complement operators may in principle be substituted. Mamdami implication (Mamdami and Assilian, 1975), which incorporates the min operator, is applied to capture overlap between conceptual entities in a single domain through time (implication in both domains is examined in Section 5).

2.3.1 Fuzzy Intersection

Fuzzy intersection of sets can be calculated using the min t-norm (Zadeh, 1965). In Eq. (2), the min t-norm operator is used to determine the intersection of two fuzzy sets *A* and *B* in an abstract universe of discourse ***U***. The membership function has a value between 0 and 1.

$$\mu \underset{\sim}{A} \cap \underset{\sim}{B}(u) = \min\,(\mu \underset{\sim}{A}\,(u), \mu \underset{\sim}{B}(u)) \mid u \in \boldsymbol{U} \tag{2}$$

Intersection using the min operator is represented diagrammatically in Fig. 4 for fuzzy ordered chord sets [5,7,2] and [2,4,6] in ***H***, indicated with an emboldened triangular line.

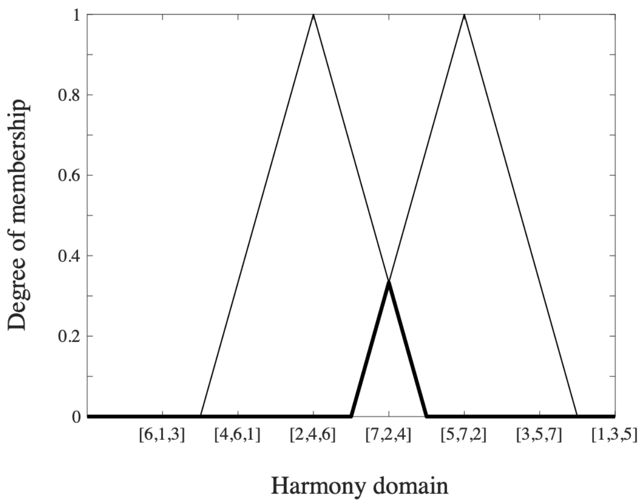


Figure 4: Fuzzy intersection of [5,7,2] and [2,4,6].

The ***H*** domain established is the key of C Ionian (C = 1, D = 2, E = 3, F = 4, G = 5, A = 6, B = 7), and the fuzzy intersection of sets [5,7,2] (G, B, D) and [2,4,6] (D, F, A) may correspond to what is traditionally described as a G major dominant ninth chord in second inversion (with added 7th and 9th tones). However, points along the intersect correspond to different semantic information. For a G major ninth chord intersect in second inversion, the bass, the pitch D of the ordered set [2,4,6] (D, F, A), is strongly salient, as is the root pitch G of the ordered set [5,7,2] (G, B, D). This amounts to an admixture of these qualities that results in a low overall membership function on the intersect. Note that this intersect accords with musical intuition: in the ***H*** domain, low membership should naturally be given to intersects where the overlap is minimal, because there is no chord category suggested in particular. By contrast, maximal overlap would be a higher membership, because a single chord category would be more strongly inferred. The partitioning structures of ***H*** and ***G*** integrate several variables for perpetual categorisation. The ***H*** domain synthesises *intervallic content*, *inversion*, *spacing* and *doubling*, and the ***G*** domain combines *instantiated events*, *relative textural density*, *proximity* and *serial parallelism* (Sections 3 and 4, respectively).

2.3.2 Fuzzy Union

The union of two fuzzy sets *A* and *B* in an abstract universe of discourse ***U*** can be calculated using any t-conorm (/S-norm). The max operator (Zadeh, 1965) in Eq. (3) is used in FRMP, yielding a membership function with a value between 0 and 1.

$$\mu \underset{\sim}{A} \cup \underset{\sim}{B}\,(u) = \max\,(\mu \underset{\sim}{A}(u), \mu \underset{\sim}{B}(u)) \mid u \in \boldsymbol{U} \tag{3}$$

The max operator is demonstrated in Fig. 5 for the fuzzy sets [5,7,2] and [7,2,4] in ***H***. The emboldened line shows the fuzzy union of these sets, and is read as *either* [5,7,2] *or* [7,2,4]. Union in ***H*** explains the notion of harmonic functionality in traditional harmonic theory (Riemann, 1905), where two or more chords can substitute for each other in the context of an overall tonic key. Functionally equivalent sets have a similar role on account of their (broadly) similar perceptual uniparametric congruence (UC). In Fig. 5, a G major chord [G, B, D] / [5,7,2] and a B diminished chord [B, D, F] / [7,2,4] fuzzily substitute for each other because they have a similar functional role in the tonal (and modal) structure, corresponding to a similar position in the partition structure. Thus, union provides a tool to categorise (and explain) the perceptual flexibility of chord set substitutes. The max operator is also important for the assembly of fuzzy relations, expanding a relation into a larger support set on a product space, which is a focus of Section 5.

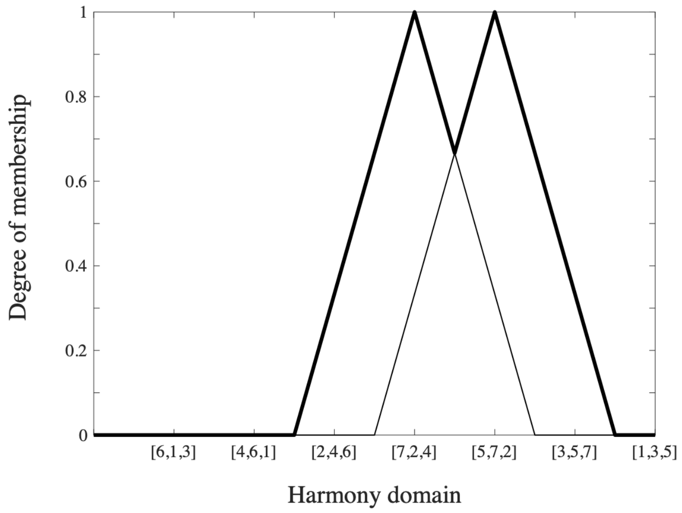


Figure 5: Fuzzy union of [5,7,2] and [7,2,4].

2.3.3 Fuzzy Complement and Negation

Similar to crisp complement, fuzzy complement (corresponding with fuzzy negation) expresses the (approximate) converse elements of a given set in a domain. In a universe of discourse ***U***, a fuzzy set, say *X*, has elements (u) with a membership function $\mu_X(u)$ between 0 and 1. The membership function of the fuzzy complement of *X* is the fuzzy converse function in the domain, where elements have a value between 0 and 1, shown in Eq. (4) (Zadeh, 1965).

$$\overline{\underset{\sim}{X}} = 1 - \mu_{\underset{\sim}{X}}(u) \tag{4}$$

The fuzzy complement of the ordered chord set [5,7,2] (G, B, D), in an ***H*** domain set to C Ionian, is shown in Fig. 6 by an emboldened line. In this case, the fuzzy complement should be understood as comprising elements that are *not* [5,7,2]. Fuzzy complement is notably distinct to crisp complement, because as shown in Fig. 6, it is possible for a single point on a fuzzy set membership function to have the same or similar degree of membership as its fuzzy complement (because the functions overlap). This conflicts with crisp set theory and classical logic, since crisp sets have the *exact* converse elements – and thus the exact converse characteristic function – as their crisp complements. The fuzzy set theory notion of complement, or its equivalent in logic, negation, is important for several aspects of music perception, such as for depicting graded harmonic change. Indeed, harmonic change at a time-point usually negates a prior harmonic entity by some degree, rather than absolutely. Also, since there is fuzzy harmonic change, there must be fuzzy harmonic rhythm, which is the rhythm (ratio of time durations) of fuzzy harmonic change, among other graded harmonic and grouping concepts.

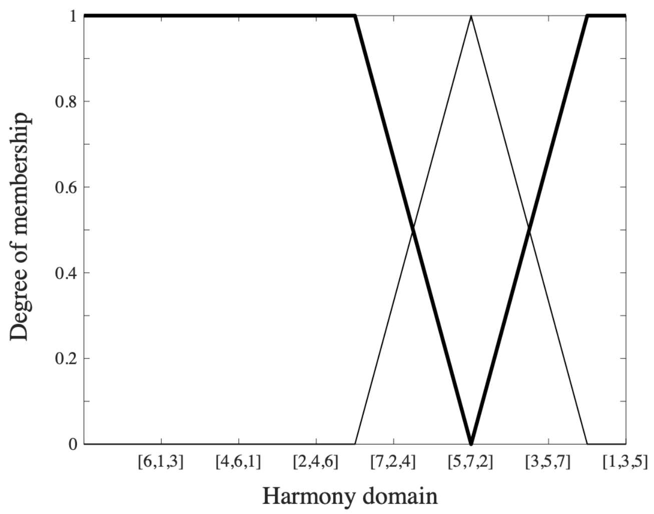


Figure 6: Fuzzy complement of [5,7,2].

2.3.4 Fuzzy Implication

As indicated above, implication in music does not correspond with the material conditional of classical logic, because in music one term does not entail another. Rather, fuzzy implication is characterised using expressions of coupling or boundedness between terms, where a graded truth value is differentially preserved through time, involving the notion of graded conceptual overlap. Music perception constructs fuzzy-logical strings of graded truth-functionality, using fuzzy implication. The standard Mamdami implication relation R_M (Mamdami and Assilian, 1975) is elected, leveraging the min operator to connect fuzzy sets, say X and Y, in a single domain, shown in Eq. (5). Implication explains various types of concept assembly and classification, such as how a series of pitches are categorised as a unified chord through time by virtue of their support for a common root, discussed in the following section.

$$R_M = \mu_{\underset{\sim}{X}} \rightarrow \mu_{\underset{\sim}{Y}}(x, y) = \min\,[\mu_{\underset{\sim}{X}}(x), \mu_{\underset{\sim}{Y}}(y)] \tag{5}$$

3 Harmony Domain

Harmonic root support and *generalised harmonic root support* determine the perception of harmonic sets and partition structure of $\boldsymbol{H}$, respectively. The notion of root support was broached in the early-eighteenth century with the theory of fundamental bass, where a single pitch, the root, provides a means to characterise a chord as a unified object that retains its identity in inversion (Rameau, 1722). Root support has been modelled formally more recently as an analogue of the pitch-perceptual process of categorising a complex tone in terms of its fundamental frequency or virtual frequency, where low-integer multiples of a fundamental frequency are defining harmonics (Terhardt, 1979, 1982; Parncutt, 1988, 1997, 2011; Milne, 2013). The four central constraints on root support are *intervallic content*, *inversion*, *spacing doubling*. While harmonic root support is important for categorising individual fuzzy chord sets, generalised harmonic root support is shown to be the basis for constructing $\boldsymbol{H}$. Generalised root support over n harmonic sets in $\boldsymbol{H}$ involves maximising the abstract hierarchical arrangement of thirds (M3/m3), perfect fifths (P5) and perfect unisons (/octaves) (P1) over the root of a key chord to determine a tonal hierarchy. It is assumed that the constraints act to construct sets and domains in working memory in real-time perception of music.

3.1 Harmonic Root Support

In Parncutt (1988), a harmony categorisation algorithm is presented where intervals of a chord set are weighted based on degree of root support, leveraging an analogy with pitch perception in a process termed sub-harmonic matching. Intervals that support a particular sub-harmonic root are called root-supporters and intervals that detract from a root are root-detractors: the unison (P1)/perfect octave (P8), perfect fifth (P5) and major third (M3) are strong root-supporters; the minor seventh (m7) and major second/ninth (M2/M9) are weak root-supporters; the minor second (m2), perfect fourth (P4), tritone (TT), minor sixth (m6), major sixth (M6) and major seventh (M7) are root-detractors; and the minor third (m3) is thought to be inert (Parncutt, 1997). The Parncutt algorithm determines a score of root ambiguity for a given chord set, such as ascribing low root ambiguity to common major and minor chords, and moderate root ambiguity to diminished, half-diminished and augmented chords (Parncutt, 1988). These calculations accord with intuition and empirical evidence: most listeners find common major and minor triadic chords to be more stable harmonic entities than other types of chords (Krumhansl, 1990; Justus & Bharucha, 2002; Huron, 2006). Fig. 7 shows root-supporting and root-detracting harmonics and intervals on the harmonic series. Pitch intervals that correspond to low-integer harmonic ratios occur early in the series (see Bernstein, 1976). Note that the root-support algorithm effectively models innate harmonic perception, since, in principle, knowledge of root support does not require learning or enculturation (corresponding with the theorisations and findings of Lerdahl, 2001; Gill and Purves, 2009; Milne, 2013; Savage et al., 2015; although, cf. Gjerdingen, 1988; Krumhansl, 1990; Justus & Bharucha, 2002; Bharucha, 2009; Huron, 2006; Patel, 2008). The Parncutt (1988) algorithm accords with cross-cultural expectations: the most common scale sets and harmonic lexicons found in cultures across the world use P1/P8, P5 and M3/m3 above a key root (Gill and Purves, 2009; Savage et al., 2015), suggesting a universal preference for perceptual or cognitive representations based on harmonic root support.

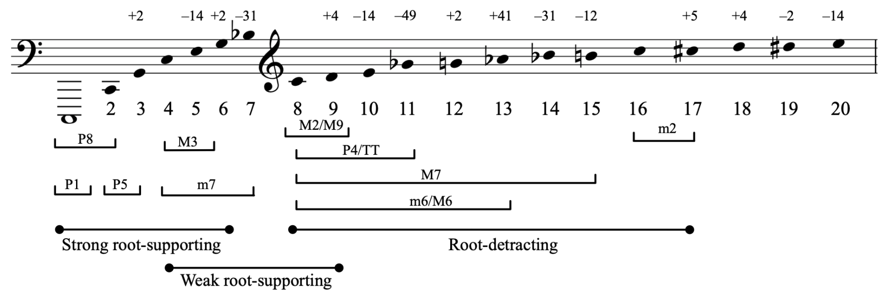


Figure 7: Harmonic series denoting strong and weak root-supporting harmonics/intervals and root-detracting harmonics/intervals (+ and – indicate differences in cents with equal temperament).

The root support of a key chord, i.e., [1,3,5], is generalised over *n* fuzzy sets to form ***H*** domains in the construction of tonal hierarchies. *Generalised harmonic root support* therefore regulates degree of UC over the tonal hierarchy. Stable UC decreases over the domain partition structure incrementally with each set (from right to left in Fig. 2). The process for determining root support and generalised root support can be shown using the concept of an *interval inclusion hierarchy*, where P1, P5, and M3/m3 ideally observe an approximate constituent hierarchy (Longuet-Higgins, 1962a, 1962b; Balzano, 1982; Lerdahl, 2001; Mazzola et al., 2002; Rawbone, 2021). The inclusion hierarchy comprises constituent pitch series in a single dimension (***H***), involving root P1 intervals, a series of m2/M2, a series of m3/M3 and a series of P5 (Fig. 8). Hierarchical inclusion acts both on individual chord sets and complete ***H*** domains, although the causative orientation of the hierarchy differs with respect to each. A chord set is constructed based on the preference for hierarchical inclusion of pitches that correspond to the spectral harmonic content of a single chord set root: M3/m3 ⊂ P5 ⊂ P1. By contrast, since root support generalises over a number of roots in ***H*** to determine tonal structure, the converse orientation is required: hierarchical inclusion of pitches corresponds to *n* roots (P1) over an abstract ***H*** superset: M3/m3 ⊃ P5 ⊃ P1. That is, to construct a single chord set (with a single root), an orientation is required where P5 and M3/m3 (in this order of preference) support a single P1 root (harmonic root support). To construct *n* fuzzy chord sets in an ***H*** superset to abstractly determine tonal structure, multiple P1 roots must be invoked by M3/m3 and P5 intervals, in this order of preference (generalised harmonic root support). A further condition is assumed to define a particular tonic: the abstract *n* roots of ***H*** must be configured to one particular sub-harmonic tone at the highest levels of structure through interaction with ***G***, involving MC relations between the domains (Section 5).

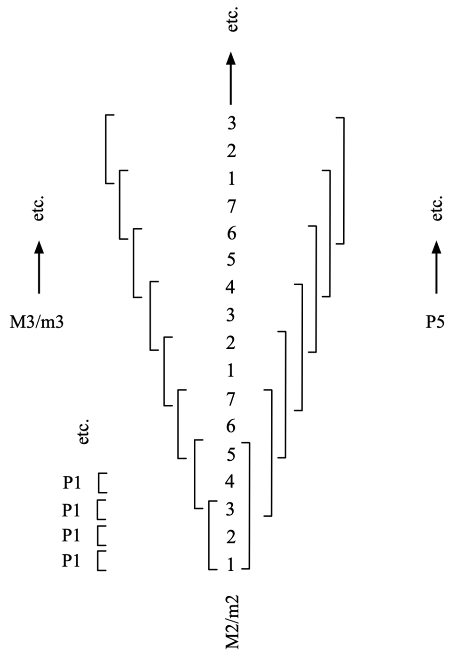


Figure 8: Schematic diagram of superimposed interval sets in a single dimension (inclusion hierarchy).

In contrast to the proposed weak root-support properties of M2, espoused in Parncutt (1988), any M2/m2 interval set, such as a standard scale set (e.g., pentatonic or heptatonic), is here thought to be redundant (/inert) in the perceptual construction of ***H***. Indeed, M2/m2 intervals seem to emerge through supervenience on higher-level interval sets in the inclusion hierarchy (although, cf. Balzano, 1982; Pearce, 2016). Accordingly, where a tritone (T) appears in a M2/m2 hierarchy in ***H*** (equivalent to two stacked m3 intervals), or in a scale set (e.g., between the seventh and fourth degrees of major scales), it emerges only as a limitation in maximising *n* roots of *n* chord sets in the inclusion hierarchy. Similar to chord sets, all interval sets in ***H*** are incrementally harmonically root-detracting (increasingly non-stable UC) with successive steps away from the key chord root. Since this harmonic picture may be captured by a single complex mathematical function (in one dimension), spatial models (with multidimensional mappings), such as pitch-space geometries or twisted tori – representing pitch, harmonic, or tonal relations – are questionable in principle (as noted above), since there is arguably no basis by which to invoke more than one dimension for harmonic perception (cf. Lerdahl, 2001; Tymoczko, 2012). It is possible, nevertheless, that spatial models are verifiable properties of higher-level cognition, but the validity of such remains an open question, requiring further theoretical support and empirical testing. In any case, spatial models seem not to characterise the perceptual sphere, which is the focus of the present theory.

3.2 Constraints on Harmonic Root Support

It has been posited that there are four central constraints on harmonic root support used for determining harmonic sets: *intervallic content*, *inversion*, *spacing* and *doubling* (based on Parncutt, 1988). The significance of intervallic content for root support has been outlined above. The constraint for inversion concerns the increased root support when root support intervals are lower relative to other tones in a harmonic set. The spacing constraint refers to the greater root support with larger relative separation of pitches. The doubling constraint, such as through unison or octave duplication, relates to increased root support when doubling root-support intervals, and conversely, the decreased root support when doubling root-detracting intervals. It is proposed that the membership functions of sets and the partition structure of ***H*** automatically encapsulate these constraints. However, a formal algorithm is not here provided which, given digital music parameters as input (e.g., MIDI), synthesises the constraints in terms of a specific mapping to the partition structure and provides membership values as outputs. At present, the process is determined manually. Fig. 9 shows two fuzzy intersections with an ordered fuzzy set [1,3,5], which is the key chord of the domain. The intersect peaking at λ_1 (in bold) involves the ordered fuzzy sets [1,3,5] and [3,5,7], and the intersect peaking at λ_2 (in bold) involves the ordered fuzzy sets [1,3,5] and [5,7,2]. The intersects are calculated using the min operator (Eq. (2)). Importantly, the intersections at λ_1 ($\mu = .66$) and λ_2 ($\mu = .33$) mark first and second inversion chord sets, respectively, but also mark chords a third and fifth higher than the key chord set, respectively, since the combined interaction of the constraints (*inter alia*) dictate mapping to the partition structure (although, as noted, a mechanism has not yet been formalised that prescribes this interaction precisely).

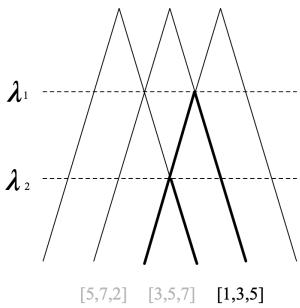


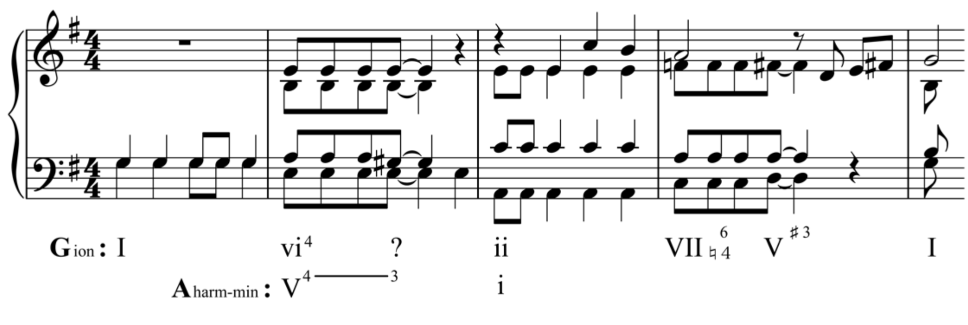
Figure 9: Partitioning in the ***H*** domain synthesising constraints intervallic content, inversion, spacing and doubling.

A synthesis of the four constraints is not possible using crisp set theory. Chord classification with crisp sets can only capture a single variable, and all four constraints are required to determine sets. A single variable, such as intervallic content, is insufficient for determining root and classifying harmony. It is undesirable to represent chord sets solely in terms of intervallic content (or inversion, say) because the four constraints are interconnected and graded. For example, with ***H*** configured to C Ionian, some λ_2 intersection of [1,3,5] and [5,7,2] may be both a second inversion tonic chord, by degree, and also a root position dominant chord with added tones, by degree. If the ordered intersection is a second inversion tonic chord set, i.e., [5,7,1,2,3], the more dissonant added tones, i.e., subset [7,2], would presumably be less perceptually salient than the consonant tones, an effect which can be achieved through varying the constraints of spacing or doubling. The upshot is that the required interaction of all four constraints necessitates fuzzy membership functions in a fuzzy partition structure.

The triangular partition structure of ***H*** also allows greater integration of intervallic content than crisp set functions, since ordered fuzzy chord sets can be extended by adding tones from adjacent sets. Using the min t-norm operator (Eq. (2)), the 7th or 9th above any chord can be appended to a main set by intersecting the main set with an adjacent set. Such added tones are either weak root-supporters or root-detractors. In Fig. 9, a seventh chord can be made by appending the [7] to the key chord [1,3,5], applying the min operator to aggregate the key chord set [1,3,5] and the adjacent set [3,5,7]. This may produce such an ordered intersectional set as [1,3,5,7], termed a chord I with an added seventh. The [1,3,5,7] ordered intersect has a maximum membership presumably reaching just short of the λ_1 level ($\mu = .66$), assuming inversion, intervallic content, spacing and doubling are minimally affected, that is, where the added seventh note has a limited effect on overall intersect membership. However, the intersection of [1,3,5] and [3,5,7] can also result in various other fuzzy classifications, depending on the interacting strengths of the four constraints (including the particular root-support qualities of the added tone(s)). For instance, the same min operation on chord I [1,3,5] with chord III [3,5,7] can also generate the ordered intersectional set [3,5,7,1]. This latter ordered intersect, *ceteris paribus*, has a lower membership (closer to the λ_2 level ($\mu = .33$), and may more strongly suggest chord III than chord I, even though the added tone above the chord bass, i.e., [1], has weak root support for chord III and strong root support for chord I. The increased root support for chord III is based on the greater significance of inversion ([3] in bass) over interval content (the "added tone" [1]) in this case, where the ordered intersect [3,5,7,1] would most strongly support the lowest tone.

A limitation with t-norms is that they cannot provide an account of 11th or 13th added chord tones. Non-linear membership functions may feasibly account for such extensions, because their curved functions can extend across a greater number of sets. However, since 11th and 13th intervals are always root-detractors, it is questionable whether such should legitimately be incorporated into harmonic representations, because these intervals by definition do not provide harmonic root support – they are de facto non-harmonic. The underlying question here is whether it is valid for perception to categorise tones into chord sets when those tones detract from the essence of those sets, because such a framework contradicts the principle of root support as a basis of harmonic categorisation. However, while integral, these questions of representation concerning the use of non-linear functions are beyond the scope of the present thesis, and so must be left for future work to explore. A further significant point is that it should be axiomatic that in the process of constructing ***H*** sets, any and all tones can be abstracted away from their crisp transduced percepts to determine a fuzzy harmonic set. For example, a crisp singleton chroma set, say {1}, *ceteris paribus*, is interpreted as a harmonic entity that generates strong root support for the fuzzy ordered set [1,3,5]. Accordingly, {1} may produce a near-maximal membership value for the [1,3,5] set, above the λ_1 level in Fig. 9, although may or may not reach the possible maximum (i.e., $\mu = 1$), because its "missing" intervallic content, i.e., [3,5], could lower net root support.

The present framework can offer only an indirect explanation of chromaticism and modulation, although as supplemental theorisation it is nevertheless significant. Modulation (/tonal reconfiguration/tonicisation) occurs when there is a change of key chord set in ***H***, where the domain is reconfigured to a novel chord in a novel tonal hierarchy to satisfy generalised root support. Ex. 2, from "I Get Around" (1964) by The Beach Boys, illustrates representation of generalised root support, hierarchical inclusion of intervals, inversion, added chord tones and modulation. As discussed, set membership is approximated based on intuitive manual integration of the four constraints, it is not calculated formally. The harmonic progression in Ex. 2, formulated in Fig. 10, involves modulation, or tonicisation, moving from G Ionian/Mixolydian {G, A, B, C, D, E, F, F♯} to A harmonic minor {A, B, C, D, E, F, G♯} in bar 2^3. The tonicisation uses a pivot (E tone) in bar 2^1, which is a tone or chord common to two keys (***H*** domains) that acts as a point of direct connection between them (Piston, 1941). Since the G♯ tone of the E major chord set [E, G♯, B] in bar 2^3 is a root-detractor of the key chord set [G, B, D], contradicting the generalised root support of the G Ionian/Mixolydian domain, a novel domain (novel tonality), A minor harmonic, is initiated to avoid paradox. The E major fuzzy chord set [E, G♯, B], provides strong generalised root support for the A harmonic minor domain, because the lowest tone E of [E, G♯, B] is a perfect fifth (P5) above the root tone (A) of the key chord set [A, C, E]. The process of generalised root support is determined by the inclusion hierarchy for ***H*** domains set out above, i.e., M3/m3 ⊃ P5 ⊃ P1/P8 (Fig. 8), although the individuation of harmonic and tonal key chord roots also requires interconnection with domain ***G*** (Section 5).



Example 2: Harmonic analysis of "I Get Around", modulating between
G Ionian/Mixolydian and A harmonic minor.

When the domain switches back to G Ionian/Mixolydian in bar 4^1, it uses an intersectional pivot chord that combines [F, A, C] and [C, E, G] sets, resulting in the ordered intersect [C, E, F, G, A], where the subset [E, G] is implicit in this instance, meaning not sounded. (However, as pointed out above, many different points may be elected along an intersect function, corresponding to the variable action of constraints.) The particular intersectional point in question has a low membership value, presumably located below the λ_2 level (see Fig. 9 and Fig. 10), falling outside the main chord categories and thus being ambiguous with respect to root. This illustrates the symbiotic connection between the constraints intervallic content and inversion (and to a lesser degree in this case, spacing and doubling), since while the intersectional point can be categorised as an F major chord set in second inversion (i.e., chord VII in G Mixolydian) owing to its intervallic content, its bass note also supports a C major chord set (i.e., chord IV in root position in G Ionian/Mixolydian). Broadly, the min operator enables synthesis of the constraints, showing fidelity to the rough and tumble of the interactive and graded harmonic landscape. It can be concluded from Ex. 2 that holding a fuzzy set in perception based on root support enables a coherent representation of the graded external musical landscape. This model of perception may be contrasted with idealised generative grammars, where the internal principles of interaction are fixed and so do not flexibly incorporate such harmonic ambiguities. Crisp theoretical categories do not mesh with graded musical structure and fuzzy perceptual processes.

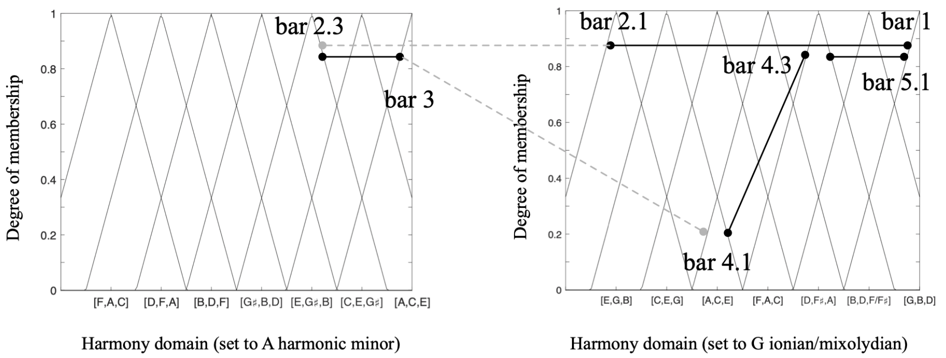


Figure 10: Approximate path through harmonic set functions in
G Ionian/Mixolydian and A harmonic minor ***H*** domains ("I Get Around").

In summary, it has been shown that fuzzy sets in ***H*** enable representation of the graded harmonic and tonal landscape. The sets and partition structure of ***H*** are based on root support and generalised root support, respectively, and determined using the constraints outlined.

4 Grouping Domain

Grouping perception involves mapping fuzzy time-regular grouping sets to time-slices in musical structure. It is proposed that *time regularity* and *generalised time regularity* are innate principles for constructing the sets and partition structure of grouping domain ***G***, respectively. The constraints on time regularity are *instantiation*, *relative textural density*, *proximity* and *serial parallelism*. The generalisation of time regularity over a hierarchy of *n* grouping sets determines the partition structure of ***G***. It is assumed that the constraints are synthesised in working memory for the construction of grouping sets and domains during real-time listening.

4.1 Constraints on Time-Regular Grouping

The initial processing of sound waves in the cochlea and auditory nerve, prior to absorption of information into the perceptual music faculty, involves transduction: complex harmonics are converted into *pitch percepts* based on support for a fundamental or virtual frequency; also, pitch events through time are converted into low-level *time-regular units*. The regular units are the beat or tactus levels, formed through bottom-up entrainment. Thus, the only sensory information represented in the music faculty is pitch and low-level time-regular percepts (Rawbone, 2021). For the ***G*** domain, low-level regular durations of time are the only basic incorrigible percepts. Many perceptual studies and computational models support the idea that representation of beat or tactus is a fundamental and automatic pre-conceptual process (Rosenthal, 1992; Todd, 1996; London, 2004; Todd and Lee, 2015; Rajendran et al., 2018; Rajendran et al., 2020). This is backed up by principles of audio scene analysis and gestalt psychology (cf. Bregman, 1990), and parallels the notion of automatic pre-conceptual processing in vision (Pylyshyn, 2001; Fodor, 2008). According with the principle of compositionality, time-regular basic percepts form the "building blocks" from which low-, mid- and high-level concepts are constructed (Cone, 1968; Benjamin, 1984; Lester, 1986; Rawbone, 2021). As argued above, while basic percepts are un-revisable, low-level concepts (harmony and grouping), mid-level concepts (e.g., harmonic rhythm) and high-level concepts (e.g., tonal and metrical structure) can be reinterpreted during the emergent unfolding of events through time.

Time-regular grouping encompasses the parameter "textural grouping" espoused in Rawbone (2017), which is the horizontal grouping of texture based on vertically corresponding event onsets in serial pitch streams. However, with textural grouping, only highly regular (highly stable UC) groups in a texture were theorised to inform metrical structure. Time regularity, by contrast, acts across slices of an individual pitch stream and combined textures, so that a grouping level set can be varied and graded at particular time points. Time regularity is informed by the four constraints and results in membership values that synthesise overall interaction. A formal model of this interaction is not presented here but determined manually. (The integration of the constraints – instantiation, relative textural density, proximity, and serial parallelism (*a–d* below) – for time-regular sets in ***G*** will be demonstrated manually in Ex. 3, Section 4.2.) The construction of regular level sets in ***G*** is an iterative process, which sometimes incorporates information from ***H***. The constraints are combined in working memory in real-time, and so level sets generated must involve relatively short strings of ***G*** elements (time-regular units) owing to intrinsic limitations of this type of memory (Deutsch, 2013a). Accordingly, terms within level sets likely involve < 10 serial elements. Of particular note, constraints *b* and *c* necessarily require repeated iterations to generate level sets > 2 serial elements, because they utilise binary relations.

The first constraint, *instantiation* (constraint *a*), acts on both monophonic and polyphonic textures, coining events on a single level that are, by degree, a multiple of (or equivalent to) units of the lowest level set (the beat or tactus level). Events in a single level are simply fixed as sets by virtue of being articulated at any time-point(s) (based on Lerdahl and Jackendoff, 1983; Temperley, 2001). A further type of instantiation event in ***G*** occurs through harmonic change, where negation and implication information is incorporated from domain ***H***. This involves negation of prior harmony and re-establishment of novel harmony through implication, using Eq. (4) and Eq. (5), respectively (further discussed in Section 5). *Ceteris paribus*, there are instantiation events in a level set of ***G*** if:

> *Constraint a*. (*instantiation*): there are events in ***G*** (or information from ***H***) marking the onsets (/points of harmonic change) of level sets that are, by degree, approximate to or multiples of beat-level time-regular units.

The next constraint, *relative textural density* (constraint *b*), applies to polyphonic textures, loosely construed, categorising level sets based on the relative mass of textural elements at several time-slices (expanding on Temperley, 2001, pp. 23–54). Textural density *td* is measured as the quantity of different tones in a particular time-slice *ts*. Similarly texturally dense time-slices mark the onsets of sets at the same hierarchical level. Let $ts_1, ts_2, ts_3, \ldots. ts_n$ be a sequence of time-slices with textural densities $td_1, td_2, td_3, \ldots. td_n$, and let relative textural density of *td*s be the Cartesian product of the *td* relations: $td_1\ R\ td_2$, $td_1\ R\ td_3$, and $td_2\ R\ td_3$, …, etc. (for all $td \in TD \times TD$). *Ceteris paribus*, the maximal *td* relation *R* (max–min) comprises *td*s that mark onsets of sets at a single level. This operation is iterated upon in real-time listening.

> *Constraint b*. (*relative textural density*): Maximal *td* relation *R* (marking onsets of sets at one level) =

$$\max\min\,[\underset{\sim}{R_a}(td_1, td_2), \underset{\sim}{R_b}(td_1, td_3), \underset{\sim}{R_c}(td_2, td_3), \ldots, etc.]$$

$$\text{for all } (td_1, td_2, td_3, \ldots, td_n) \in TD \times TD$$

Proximity (constraint *c*) concerns pitch or time events in both monophonic and polyphonic music. Proximity largely follows the rules and theories of Lerdahl & Jackendoff (1983, pp. 345–346), Bregman (1990), Deutsch (1998, 2013a, 2013b), and Temperley (2001). Events of either pitch or time units that are most proximal in pitch or time form a collective series, the duration of which is whole or part of a level set. Let $e_1, e_2, e_3, \ldots e_n$, be either pitch or time events e, and let proximity be the Cartesian product of e relations: $e_1 R e_2$, $e_1 R e_3$, and $e_2 R p_3$, …, etc. ($E \times E$). *Ceteris paribus*, the relation with the maximal proximity (max–min) comprises events which form a collective series, and as stipulated, the duration of that series forms whole or part of a single level set. This operation is iterated upon in real-time listening.

> *Constraint c*. (*proximity*): Maximal proximity relation *R* (duration of whole or part of a single level set) =
>
> max min $[\underset{\sim}{R_a}(e_1, e_2), \underset{\sim}{R_b}(e_1, e_3), \underset{\sim}{R_c}(e_2, e_3), \ldots, etc.]$ for $(e_1, e_2, e_3, \ldots, e_n) \in E \times E$

Serial parallelism (constraint *d*) concerns both monophonic and polyphonic music. *Ceteris paribus*, approximately parallel serial streams of either pitch or rhythm (rhythm defined as a ratio of time durations), that is, *a*, *b*, *c*, …, *n*, or *a*: *b*: *c* …. *n*, form corresponding sets, the lengths of which are the duration of level sets (based on Lerdahl and Jackendoff, 1983, p. 51).

> *Constraint d*. (*serial parallelism*): *a*, *b*, *c*, …, *n* ≈ *a*, *b*, *c*, … , *n* (*pitch parallelism*)
>
> *a: b: c*, …, *n* ≈ *a: b: c :* , … , *n* (*rhythm parallelism*)

Proximity and parallelism of grouping are viewed as automatically-perceived properties (gestalts) in generative theories (Lerdahl and Jackendoff, 1983; Bregman, 1990; Temperley, 2001). Since these constraints interact with each other in the ***G*** domain, and with other sets and relations in ***H***, involving both bottom-up and top-down processing, such are not bone fide gestalt mechanisms. However, these are still *relatively* automatic perceptual processes and are nonetheless here considered fundamental constraints of the ***G*** domain. As noted, a formal model that integrates constraints *a–d* and delineates their interconnections with ***H*** is yet to be determined. Although, *prima facie*, it might be possible to construct ***G*** sets based on an overall relation between the four constraints, in many circumstances the interaction is more nuanced, and a feasible solution has not emerged. The process is not trivial and requires further theorising and empirical work.

4.2 Formation of Sets and Partition Structure

A fuzzy set in ***G*** is the fuzzy grouping of a duration between two time-points, involving fuzzy regularity of a hierarchical level determined by constraints *a–d*. (The relative values of grouping levels are as follows: maxima = 8; longa = 4; breve = 2; semibreve = 1; minim = .5; crotchet = .25; quaver= .125; semiquaver = .0625.) Fig. 11 shows the hierarchical basis of partition structures. The partition structure embodies generalised time regularity, where a main time-regular level *L* differentially supports time-regular events at lower levels, shown with arrows in Fig. 11. A main level *L* is two or three times the time length of a sub-level set *sl* ($L = 2 \times sl$ OR $3 \times sl$) and four or six times the time length of a sub-sub-level set *ssl* ($L = 4 \times ssl$ OR $6 \times$ ssl). In Fig. 11, the main level *L*, a longa level set [4], supports the double semibreve-level set [2] and, to a lesser degree, supports the semibreve-level set [1].

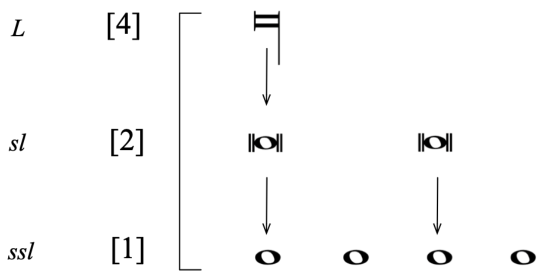


Figure 11: A fuzzy main level set L and lower-level sets, sl and ssl, in $\boldsymbol{G}$.

Generalised time regularity involves differential UC between a main level set and lower-level sets in $\boldsymbol{G}$. In Fig. 12, the main grouping level is likewise fixed at [4], which has maximally stable UC. The stable UC decreases incrementally with each set over the partition structure (from right to left), comprising a fuzzy $\boldsymbol{G}$ partition structure with intersecting sets at λ_1 (μ = .66) and λ_2 (μ = .33). Intersection involves the time-regularities of any two fuzzy sets aggregated at a single time-point (see also Fig. 3). The emboldened fuzzy intersect at λ_1 (μ = .66) involves aggregation of the longa set [4] with the breve set [2]. An example of such an ordered intersect is [4,2], where [4] is the most dominant level. The order depends on which level is most salient based on constraints *a–d*. Similarly, the emboldened intersection at λ_2 (μ = .33) involves aggregating the longa set [4] and the semibreve fuzzy set [1], which results in an ordered fuzzy intersectional set such as [4,1]; again, the order of the intersect depends on which level subset is most salient. Intersection follows the intuition that at any particular time-point it is possible to represent a fuzzy admixture of level sets in perception, and where the membership function of intersects is lower than main fuzzy categories (at λ_1 or λ_2). Ex. 6 in Section 5.5 illustrates fuzzy intersection of sets in $\boldsymbol{G}$, assigning semibreve and minim sets to particular time-points (notated half-bar time-points).

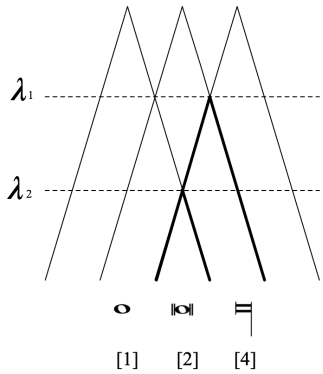
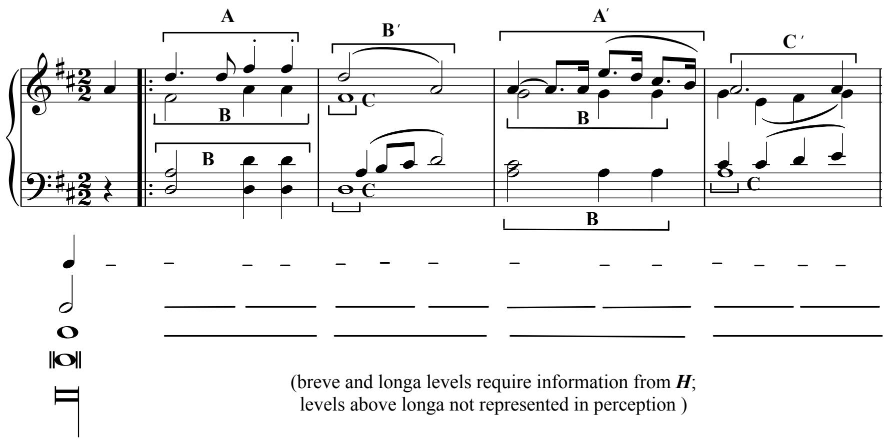


Figure 12: Fuzzy semibreve set [1], breve set [2], and
longa set [4], and their intersectional sets at λ_1 at λ_2 in $\boldsymbol{G}$.

A hierarchy of level sets in $\boldsymbol{G}$ can now be demonstrated in the context of a musical texture. Ex. 3, from Schubert's "Trout" Quintet, D. 667, iv (1819), shows a fuzzy grouping hierarchy, depicted using the line nomenclature of Temperley (2001). The level sets are not perfectly crisp, although relatively uniform, which is characteristic of homophonic music. Quaver-level sets [.125] and lower are too sparse to form significant concepts in the conceptual hierarchy. The fuzzy crotchet-level sets [.25] are the lowest established level sets, both through instantiation (constraint *a*) and relative textural density (constraint *b*), on notated beats 1, 3 and 4 of most bars. Minim-level sets [.5] are individuated by instantiation (constraint *a*) and relative textural density (constraint *b*) at every notated bar and half-bar onsets. Semibreve-level sets [1] are determined through instantiation (constraint *a*) and the approximately serial pitch and rhythm parallelisms (constraint *d*) of all the motifs A–A′, B–B′ and C–C′ at notated bar onsets. However, the fuzzy breve level sets [2] and longa level sets [4] must be individuated using information from the ***H*** domain, involving resemblance relations R_R with the ***G*** domain, examined in Section 5. Grouping levels (and metrical levels, more abstractly) above the longa level are not articulated in perception, although may be represented by indeterminate processes in cognition.



Example 3: Fuzzy grouping hierarchy in
Schubert's "Trout" Quintet, D. 667, iv (1819).

In summary, the sets and partition structure of ***G*** are constrained by time regularity and generalised time regularity, based on the constraints and theorisation presented. This formalisation provides a principled approach for determining grouping structure.

5 A Model of Fuzzy Relational Music Perception

This section presents FRMP, which concerns two main relations: fuzzy resemblance relations R_R (also termed compatibility or tolerance relations (Ross, 2010)), which involve analogical resemblance between ***H*** and ***G***; and fuzzy logical implication relations, which represent continuation through time (or overlap) in one domain (R_M) or both domains (R_I). While R_R comprise the properties of weak-reflexivity, weak-symmetry and antitransitivity to determine conceptual hierarchy, implication relations, R_M and R_I, concern the intersectional overlap of sets through time, leveraging the min operator. Relations between ***H*** and ***G*** are necessary and sufficient for the individuation and assembly of concepts, from low-level concepts, such as event groups and chords, to high-level concepts, such as tonal and metrical structures. Fuzzy relations between domains are binary, involving the interaction or interconnectedness between elements of ***H*** and ***G*** sets, expressed as Cartesian product (Zadeh, 1971; Klir & Yuan, 1995). As noted above, it is axiomatic that concepts are individuated and assembled compositionally, from basic to complex. Perception represents as many levels of abstraction as computationally possible, where higher-level sets (/concepts) are constructed from lower-level sets (/concepts), right down to basic percepts. Compositionality characterises the structure of domains and their relational interaction, permitting the generation of coherent and comprehensible higher-level concepts. It is also axiomatic that relations between ***H*** and ***G*** are informational, not absolute: sets across the disparate domains are connected by analogical resemblance, which means they have a similar level of UC between their disparate domains. Also, the domains are scalable to enable coherent interaction, where ***G*** can be protracted or contracted to cohere with ***H*** and vice versa. Unarticulated sets in-between articulated sets of domains can be omitted, providing the partition structures are otherwise preserved (since the partition structures of both domains are intrinsic and universal).

It is assumed that fuzzy relations at a single time-point are stimulated, or "fired", in the music faculty on the occasion of sense data. Fuzzy time-point relations represent simultaneous events in $\boldsymbol{H}$ and $\boldsymbol{G}$, where the external data triggers innately-encoded relational neural architecture. The Cartesian relational architecture of the biological neural network may alternatively be predetermined by innate constraints. In terms of this latter contingency, the network may be considered innate-in-effect, because neonatal learning must be rigidly constrained toward representing only information expressible as relations. Several time-point relations are aggregated through time to form R_R, R_M and R_I, which are represented on a $\boldsymbol{H} \times \boldsymbol{G}$ product space. Conceptual structure is implicated through time as a truth-functional logical form using R_M and R_I. The overlap is calculated by aggregating the relations between two or more time-point sets or relations, in one or both domains, respectively. When a concept is introduced that does not form a fuzzy implication of a previous concept – that is, where there is negation – R_R are required, permitting the construction of hierarchical structure through time. The max operator aggregates two or more time-point relations on a product space for the construction of R_R. R_R unite different level sets under a single hierarchy (in product space $\boldsymbol{H} \times \boldsymbol{G}$), where more stable UC subset relations are at higher levels than less stable UC subset relations. The relational hierarchy results in tonality and metre being interdefined by R_R. In order to isolate a concrete set within the tonal–metrical hierarchy, the dimensionality of a subset relation is reduced to one of either domains, using projection (proj), explained in Section 5.1 (Eq. (9)). The use of relations in music can be highly complex, and there are circumstances where multiple relational hierarchies are set up concurrently, termed *plural relational hierarchies* (Section 5.5).

Fig. 13 shows a schematic diagram of FRMP, which presumes iterative relation formation and concept construction. Implication relations in one domain (R_M), $\boldsymbol{H}$ or $\boldsymbol{G}$ (Eq. (5)), are negated within the same domain, using Eq. (4). Implication in both domains simultaneously (R_I) ($\boldsymbol{H}$ and $\boldsymbol{G}$) is determined by Eq. (15), and negated by substitution of R_I into Eq. (4). As noted, relations of resemblance (R_R) (Eq. (8)), are formed between domains ($\boldsymbol{H}$ and $\boldsymbol{G}$) through time. $\boldsymbol{H}$ and $\boldsymbol{G}$ subsets within R_R and R_I are concretised using projection (proj), as outlined above. All these notions rest on theorisation and equations set out in Sections 5.2–5.4. A further consideration is that perceptual relations and concepts may later be made crisp by higher-level systems in cognition, such as by consciousness, and memory and belief systems. Crisp concepts enable efficiency of representation and consolidation, achieved using various types of defuzzification methods, although high-level cognitive processing is a peripheral consideration in the discussion here.

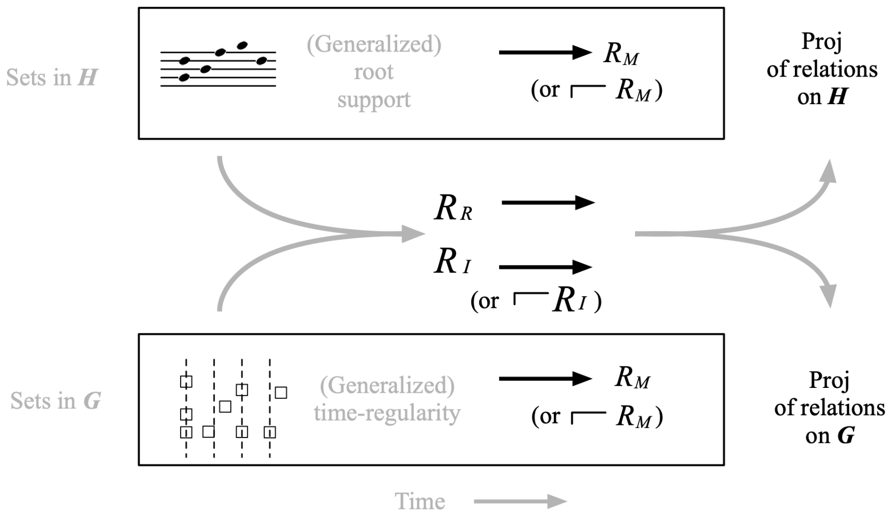


Figure 13: Schematic diagram of relation formation and concept individuation and assembly in FRMP.

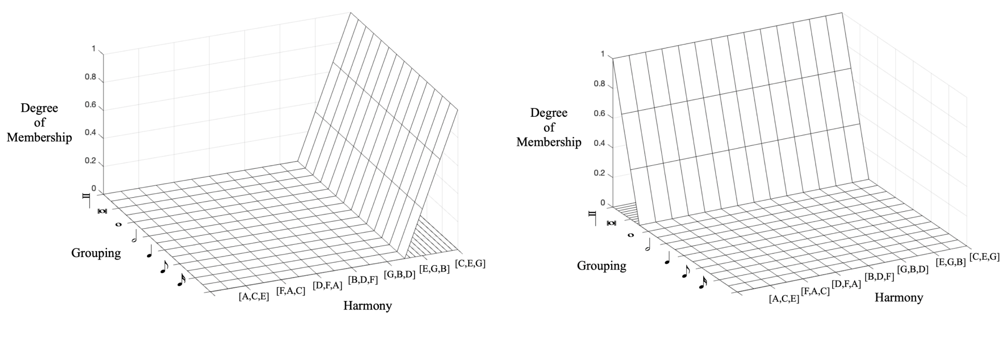
The following explication presents the mathematical components of FRMP, outlining the properties and characteristics of R_R, R_M and R_I, and indicating how these are used in perception for concept construction. A further aim is to explore plural relational hierarchies to account for contrapuntal architectures.

5.1 Constructing Time-Point Relations and Resemblance Relations

Fuzzy relations between $\boldsymbol{H}$ and $\boldsymbol{G}$ can be represented on a $\boldsymbol{H} \times \boldsymbol{G}$ Cartesian product space, involving ordered pairs of elements that are ascribed degrees of truth between 0 and 1. Relations at a single instance in time are represented in neural architecture, depicting interconnections between the domains at a time-slice, termed "time-slice relations". However, on the occasion of sense data, relations are represented concretely by a point within the time-slice of the product space. Resemblance relations R_R and implication relations in a single domain (R_M) or both domains (R_I) are formed between two or more time-point relations. In this subsection, R_R are given focus; in Section 5.4, R_M and R_I are examined. The first step in representing a single time-slice relation is to form cylindrical extensions on sets of both domains, which involves adding a new domain (/dimension) to a fuzzy set to create n + 1 domains, and which generates a mapping between them, where a set is extended over the product space (Zadeh, 1971; Klir & Yuan, 1995). Eq. (6), based on Garriga–Berga (2005), shows the cylindrical extension (ce) of a harmonic set X on the $\boldsymbol{H} \times \boldsymbol{G}$ product space. A cylindrical extension should be carried out, *mutatis mutandis*, on a set, say Y, in the $\boldsymbol{G}$ domain. Thus, a fuzzy extension of X ($h \in \boldsymbol{H}$) and a fuzzy extension of Y ($g \in \boldsymbol{G}$) are extensions over the same product space.

$$\text{ce of X on } \boldsymbol{H} \times \boldsymbol{G} = \{(h, g), \mu X(h)\} \tag{6}$$

To take a concrete example, in the key of C Ionian, a harmony set h ($h \in \boldsymbol{H}$) of chord I (C chord) can be extended over a product space, where the degree of membership for each C chord element is copied to all (h, g) with the same h, as shown in Fig. 14 (a). Likewise, a grouping set g in $\boldsymbol{G}$ ($g \in \boldsymbol{G}$) of the longa level can be extended over the same product space, where the degree of membership for each longa level set is copied to all (h, g) with the same g, shown in Fig. 14 (b).



(a) (b)

Figure 14: Cylindrical extensions of (a) harmony set (C chord) and (b) grouping set (longa level).

The next step in finding a single time-slice relation is to calculate the intersection of these extensions. A time-slice relation is the intersection of extensions P and Q on the $\boldsymbol{H} \times \boldsymbol{G}$ product space, calculated using Eq. (7) which involves the min t-norm operator.

$$\mu \underset{\sim}{P} \cap \underset{\sim}{Q}(h, g) = \min\left(\mu \underset{\sim}{P}(h, g), \mu \underset{\sim}{Q}(h, g)\right) \tag{7}$$

Fig. 15 illustrates the intersection between the C chord extension set (P) and the longa level extension set (Q), to produce a relation at a single time-slice (Eq. (7)). This time-slice relation is invoked implicitly and automatically in perception in real-time when such concrete sets coincide at an instance in time. Note that the absolute length of ***H*** concepts is not a primary factor in determining these relations, providing the elements concerned are perceptually salient. The relation in Fig. 15 represents the innate neural architecture stimulated by the general concept categories involved. A concrete relation, fired on the occasion of sense data, would be a single point in the time-slice subset of the product space, the position of which would depend on the properties of the actual elements involved. The relative degree of membership between domains classifies a time-point relation. That is, low-level membership between the domains is a weaker relation, and means that an event is accordingly less salient for higher-level relational sets (/concepts). Broadly, harmonic and grouping time-slice and time-point relations are precisely synthesised using this method of dimensional intersection. (Musical examples that show the intersectional method of synthesising the domains are considered in Section 5.3.)

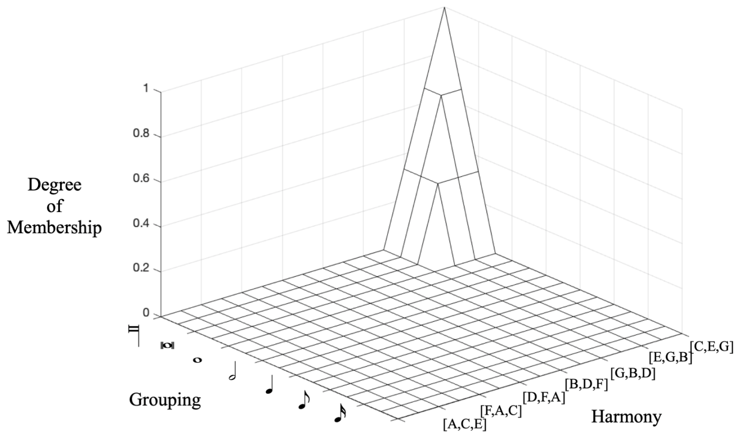


Figure 15: Time-slice relation showing intersection of a general harmony set (C chord) and a general grouping set (longa level) at a single time-slice.

The next stage in concept individuation and assembly involves the chaining together of time-slice relations to form an overall relation through time, generating R_R, R_M, or R_I. (Section 5.4 focuses on R_M and R_I.) To form an overall R_R through time, r time-slice fuzzy relations are aggregated in a fuzzy union using the max operator, shown in Eq. (8).

$$\mu \underset{\sim}{R} \cup \underset{\sim}{S} \cup, \ldots, r\ (h, g) = max\ (\mu \underset{\sim}{R}\ (h, g), \mu \underset{\sim}{S}\ (h, g), \ldots, r) \qquad (8)$$

A diagrammatic union of time-slice relations *C-major–longa R* and *G-major–breve S* is shown in Fig. 16, forming a *C major–longa—G major–breve* R_R, which represents the general neural architecture involved. An actual concrete R_R would involve four points in the relational subset, for each concrete element of the four categories. The calculations for constructing a *G-major–breve*, involving cylindrical extension and intersection (min) set out above, do not need to be repeated. While the properties of R_R are nuanced and given attention in Section 5.2, it may already be apparent that the R_R in Fig. 16 has the properties of reflexivity, symmetry and anti-transitivity in the occupied subset of the product space. Also, it can be observed that the domains are interdefined by the R_R, since R_R determine tonal–metrical hierarchies. Fig. 16 depicts a tonal–metrical hierarchy where the subset relation [C, E, G] *R* [4] are the higher levels, and [G, B, D] *R* [1] are the lower levels. Indeed, the R_R depicts a tonal–metrical hierarchy within which a C major chord set and a longa level set are the highest UC sets, which give rise to abstract tonal–metrical structure. Section 5.3 outlines musical examples that generate conceptual hierarchies using R_R.

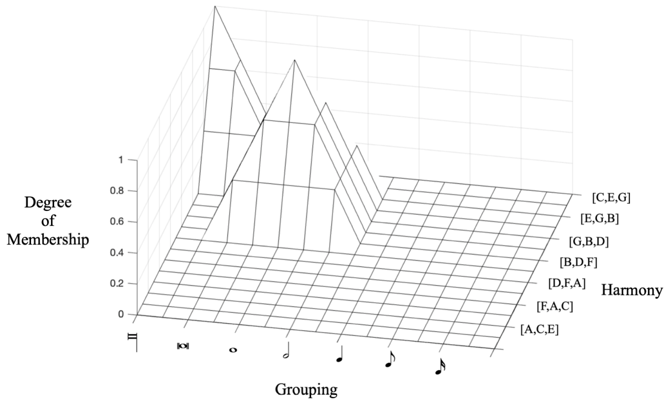


Figure 16: *C-chord–longa – G-chord–semibreve* R_R.

While abstract relations are vital for perceptual connections, the concretisation of concepts is also necessary for perception to construct concept hierarchies. Concretisation might also be used by general cognitive systems, such as consciousness or memory and emotion systems, facilitating the consolidation of perceptual information (Fodor, 1983). Thus, it is useful to incorporate the broadly converse operation of cylindrical extension, projection, to turn a relation back into a single-dimension concrete fuzzy set, reducing the dimensionality from two domains to a single domain. The projection (proj) of a subset relation (h,g) of the $\boldsymbol{H} \times \boldsymbol{G}$ product space on $\boldsymbol{H}$ is shown in Eq. (9). This equation can be used, *mutatis mutandis*, for projections of a subset relation (h,g) on $\boldsymbol{G}$.

$$\text{proj of } \boldsymbol{H} \times \boldsymbol{G} \text{ on } \boldsymbol{H} = \left\{h, \max_{g} \left(\mu_R(h,g)\right)\right\} \tag{9}$$

An example of projection on $\boldsymbol{H}$ is shown diagrammatically in Fig. 17, where the *G-chord–semibreve* relation is transformed back into a *G-chord* fuzzy set*,* shown with shadow. Note that projection maximises only those elements h by virtue of their interaction with $\boldsymbol{G}$ in the relational subset (h,g) of the product space. (A musical example of projection is given in Section 5.4.)

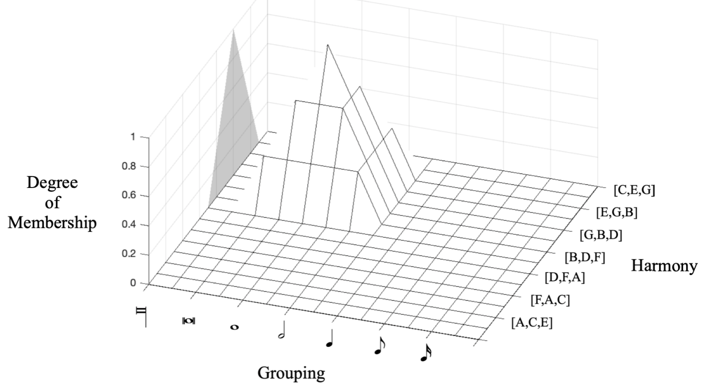
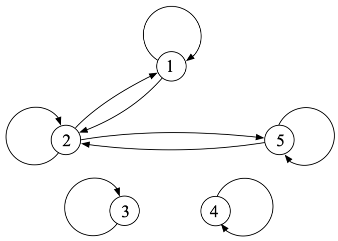


Figure 17: Projection (shadow) of *G-chord–semibreve* relation onto domain $\boldsymbol{H}$.

5.2 Properties of Fuzzy Resemblance Relations

In traditional fuzzy set theory, binary resemblance relations (R_R) are conventionally examined on a single set, and have the following properties: reflexivity, which is an equality that holds between an element and itself; symmetry, a mirroring of element connections across a relation; and non-/anti-transitivity, a lack of connection between non-adjacent elements in a chain. It is useful to initially explore these properties on arbitrary crisp sets. It can be seen that a crisp R_R is formed on the vertices and matrix diagrams in Fig. 18 owing to its reflexive, symmetrical and nontransitive properties. The R_R is crisply reflexive because in the vertices diagram (Fig. 18 (a)) every element is connected to itself, and in the diagonal subset of the matrix, all corresponding elements form connections with values of 1 (Fig. 18 (b)). The R_R is symmetrical since in the matrix and vertices diagrams there is a mirroring of elements in the relational structure. It is also nontransitive, because the three-element sequence, 1, 2 and 5 is an inequality.



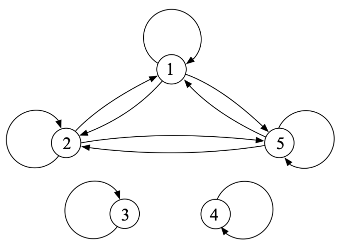
(a)

$$\boldsymbol{R}_R = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

(b)

Figure 18: (a) Crisp five-vertices diagram and
(b) crisp matrix, both showing a R_R (in Ross, 2010, p. 64).

A crisp R_R can be contrasted with a crisp equivalence relation, R_E, since the latter is reflexive, symmetrical and *transitive*. An example of an R_E is shown in a five-vertex graph (Fig. 19 (a)) and also on a relational matrix (Fig. 19 (b)). In this R_E, elements 1, 2 and 5 form a chain that is a transitive equality. In general, R_E are not useful for music perception, because they do not permit elements within a relation to be differentiated over the product space (explained further below).



(a)

$$\boldsymbol{R}_E = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

(b)

Figure 19: (a) Crisp five-vertices fuzzy diagram and
(b) crisp matrix showing an R_E (in Ross, 2010, p. 65).

Crisp matrices and vertices diagrams can easily be converted to fuzzy relations by adding membership values between 0 and 1 in the matrix and in the connecting arrows of the vertices graph. Although binary fuzzy R_R are traditionally demonstrated on a single fuzzy set, i.e., $\boldsymbol{R}\ (\boldsymbol{X} \times \boldsymbol{X})$, they are here expressed on an $\boldsymbol{H} \times \boldsymbol{G}$ Cartesian product space to simulate analogical relations. Standard definitions of fuzzy reflexivity, symmetry and antitransitivity are shown in Eq. (10), Eq. (11) and Eq. (12), respectively (Klir & Yuan, 1995).

$$\mu_{\underset{\sim}{R}}(h,g) = 1 \text{ for all } h,g \in \boldsymbol{H} \times \boldsymbol{G} \ (\textit{reflexivity}) \tag{10}$$

$$\mu_{\underset{\sim}{R}}(h,g) = \mu_{\underset{\sim}{R}}(g,h) \text{ for all } h,g \in \boldsymbol{H} \times \boldsymbol{G} \ (\textit{symmetry}) \tag{11}$$

$$\mu_{\underset{\sim}{R}}(x,z) < \max_{y \in Y} \min\,[\mu_{\underset{\sim}{R}}(x,y), \mu_{\underset{\sim}{R}}(y,z)] \text{ for all } x,z \in \boldsymbol{H} \times \boldsymbol{G} \ (\textit{antitransitivity}) \tag{12}$$

In standard fuzzy set theory, a fuzzy binary relation $\boldsymbol{R}$ ($\boldsymbol{H} \times \boldsymbol{G}$) is reflexive if it satisfies Eq. (10). If Eq. (10) does not hold for some members of $\boldsymbol{H} \times \boldsymbol{G}$, then the relation is irreflexive; if Eq. (10) does not hold for all members of $\boldsymbol{H} \times \boldsymbol{G}$, then the relation is antireflexive (Lin & Lee, 1996). A fuzzy binary relation is symmetric if it satisfies Eq. (11); if this equality is not satisfied for some elements of the relation it is termed asymmetric; if the equality is not satisfied for the elements of the relational support then it is termed antisymmetric; if the equality is not satisfied for all h, g $\in \boldsymbol{H} \times \boldsymbol{G}$ then the relation is strictly antisymmetric. A fuzzy binary relation is antitransitive if it satisfies Eq. (12). If Eq. (12) does not hold for some members of $\boldsymbol{H} \times \boldsymbol{G}$, then the relation is nontransitive; if Eq. (12) does not hold for all members of $\boldsymbol{H} \times \boldsymbol{G}$, then the relation is transitive. The following matrices $R_{\text{a-f}}$ in Fig. 20 demonstrate a number of these relational properties, some of which are used in defining R_R.

$$R_a \begin{bmatrix} 1 & 0.8 & 0.3 \\ 0.3 & 1 & 0.6 \\ 0.4 & 0 & 1 \end{bmatrix}, R_b \begin{bmatrix} 0.3 & 1 & 0.9 \\ 0 & 0.7 & 0.2 \\ 0.5 & 0 & 0.3 \end{bmatrix}, R_c \begin{bmatrix} 1 & 0.5 & 0.7 \\ 0.5 & 0.3 & 0.1 \\ 0.7 & 0.1 & 0 \end{bmatrix},$$

$$R_d \begin{bmatrix} 1 & 0 & 0.6 \\ 0 & 0.3 & 0.8 \\ 0.5 & 0.7 & 0.5 \end{bmatrix}, R_e \begin{bmatrix} 1 & 0 & 0.6 \\ 0.1 & 0.3 & 0.8 \\ 0.5 & 0 & 0.5 \end{bmatrix}, R_f \begin{bmatrix} 0.1 & 0.5 & 0.7 \\ 0 & 1 & 0.2 \\ 0 & 0.3 & 0.2 \end{bmatrix}.$$

Figure 20: Fuzzy relational properties
in matrices $R_{\text{a-f}}$ (Lin & Lee, 1996, p. 48).

It can be seen that R_a is reflexive, R_b is antireflexive, R_c is symmetric, R_d is antisymmetric, R_e is strictly antisymmetric, and R_f is transitive. Such properties are variously combined in relations. For example, the fuzzy relation "*x* and *y* are very near" is reflexive, symmetric and antitransitive; the fuzzy relation "*x* and *y* do not look alike" is antireflexive, symmetric and anti/nontransitive; the fuzzy relation "*x* is greatly smaller than *y*" is antireflexive, strictly antisymmetric and transitive (Lin & Lee 1996, p. 48). We are presently concerned only with fuzzy resemblance relations R_R, and of those, only R_R that are reflexive, symmetrical and *antitransitive*. Only the first of the above relations, "*x* and *y* are very near" is a fuzzy resemblance relation, which can be otherwise stated as "*x* and *y* are approximately equal". As noted, R_R are distinct from R_E, since the former are nontransitive or antitransitive, whereas the latter are transitive. In general, R_E are not useful in music perception to connect $\boldsymbol{H}$ and $\boldsymbol{G}$ because antitransitivity is required to differentiate between analogically similar and non-similar elements on the product space. Transitivity connects analogically different (non-corresponding) elements, which if used in perception would mean there would be no way to distinguish analogically corresponding elements from non-corresponding elements across the product space. It should be observed, perhaps confusingly, that while crisp equality and fuzzy equality (or fuzzy equivalence) are paradigm R_E, *approximately equal is not an* R_E because approximately equal can involve multiple incremental changes across a product space that can result in significant overall inequalities over that space. Rather, approximately equal is a fuzzy R_R on account of its anti-/non-transitive property (Cock & Kerre, 2001; Cock & Kerre, 2003a, 2003b; Klawonn, 2003; Beg & Ashraf, 2010).

The standard depiction of fuzzy reflexivity in Eq. (10) is too strong. A weaker form of reflexivity, weak reflexivity (w-reflexivity), enables a broader support set in the $\boldsymbol{H} \times \boldsymbol{G}$ product space and more generalised membership functions, shown in Eq. (13) (based on Yeh, 1973; Gupta & Gupta, 1996; Chon, 2003). Eq. (13) shows that for w-reflexivity, the infimum (inf) of the transitive bound $\mu(t,t)$ between the analogically corresponding elements (x, x) between the domains of the $\boldsymbol{H} \times \boldsymbol{G}$ product space must be greater than the connections between x elements and other elements y, i.e., $\mu(x,y)$, over the whole product space. Thus, w-reflexivity permits more generalised membership for analogically reflexive relations of $\boldsymbol{H} \times \boldsymbol{G}$.

$$\mu \text{ is } \textit{w-reflexive} \text{ iff } \mu(x,x) \geq \epsilon > 0 \text{ and } \inf_{t \in X} \mu(t,t) \geq \mu(x,y)$$
$$\text{for all } x, y \in \boldsymbol{H} \times \boldsymbol{G} \text{ such that } x \neq y \; (\textit{w-reflexivity}) \tag{13}$$

Similarly, the standard notion of symmetry in Eq. (11) is too strong, and a weaker form in Eq. (14) has been introduced, i.e., weak symmetry (w-symmetry) (Chon, 2017), applied here to the $\boldsymbol{H} \times \boldsymbol{G}$ product space. Eq. (14) shows that symmetrically positioned elements in the matrix can have a generalised connection, but non-symmetrically positioned elements have a zero connection. W-symmetry permits classification of a more generalised support set for fuzzy symmetrical relations.

$$\mu \text{ is } \textit{w-symmetrical} \text{ iff } \min\,[\mu(h,g), \mu(g,h)] > 0$$
$$\text{or } \mu(h,g) = \mu(g,h) = 0 \text{ for all } h, g \text{ in } \boldsymbol{H} \times \boldsymbol{G} \; (\textit{w-symmetry}) \tag{14}$$

The standard fuzzy definition of antitransitivity is required for the present model, shown in Eq. (12), where a fuzzy binary relation $\boldsymbol{R}$ ($\boldsymbol{H} \times \boldsymbol{G}$) is antitransitive if this inequality is satisfied for all elements. Antitransitivity, which holds between elements in any chain in $\boldsymbol{H} \times \boldsymbol{G}$, is important for analogical R_R between the domains, since as highlighted, only $\boldsymbol{H}$ and $\boldsymbol{G}$ elements that have a similar degree of UC stability should be connected to each other; non-similarly stable UC elements should not be related. An idealised example of a R_R is shown across a complete $\boldsymbol{H} \times \boldsymbol{G}$ product space in Fig. 21, satisfying the defining properties of w-reflexivity, w-symmetricality and antitransitivity (in fact, it is reflexive, symmetrical and antitransitive). Such R_R connect n fuzzy sets in $\boldsymbol{H}$ with n fuzzy sets in $\boldsymbol{G}$, and, during activation in perception, would involve an aggregation of r binary relations between $\boldsymbol{H}$ and $\boldsymbol{G}$ at several time-points, a process more fully examined in the following subsection.

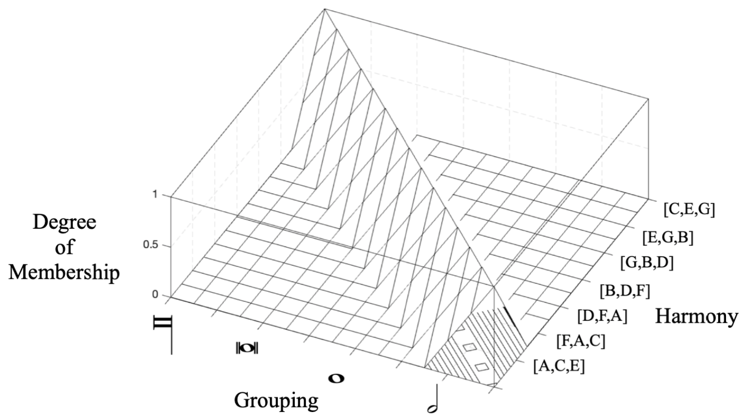


Figure 21: Idealised fuzzy resemblance relation R_R on the $\boldsymbol{H} \times \boldsymbol{G}$ product space, configured to C Ionian and the longa level.

Such a R_R over a complete $\boldsymbol{H} \times \boldsymbol{G}$ product space, as in Fig. 21, is exceptional, since more often fuzzy relations are formed as subsets of a possible product space. In Fig. 21, in the $\boldsymbol{G}$ hierarchy, the longa-, breve-, semibreve- and minim-level sets (i.e., [4], [2], [1], [0.5]), are scaled up ($\times$ 2) to match fifths steps in $\boldsymbol{H}$ (i.e., [C, E, G], [G, B, D], etc.). As discussed, scaling of a domain and omission of sets are sometimes required to construct coherent $\boldsymbol{H} \times \boldsymbol{G}$ relations, and are valid providing the partition structure – the internal scaling, set order and intersectional structure – is preserved.

5.3 Iterative Assembly of Resemblance Relations

The iterative use of R_R enables the construction of conceptual hierarchies. As discussed, relations at a single time-point are fired on the occasion of sense data (pitch and low-level time-regularity percepts). Concept construction during real-time listening involves the aggregation of time-point sets or relations to form the through-time relations R_R, R_M and R_I. Also, since concept construction is compositional, higher-level R_R, R_M and R_I are formed from concepts individuated by lower-level R_R, R_M and R_I. This section focuses on the iterative assembly of R_R, while the following section examines R_M and R_I.

The opening of Schubert's "Trout" Quintet, D. 667, iv (1819), in Ex. 3 above, involves iterative aggregation of binary R_R, using the process outlined in Section 5.2 to construct conceptual hierarchies. Let us first examine low-level aggregation, looking at the upbeat before bar 1 and bar 1^1. The domains of the matrix must be scaled to enable sets to correspond across the respective domains, a process carried out automatically in perception. In this case, the $\boldsymbol{G}$ domain is scaled-up to correspond to the $\boldsymbol{H}$ domain. Thus, $\boldsymbol{H}$ sets are partitioned into intersects of 0.33 and 0.67, and $\boldsymbol{G}$ sets are scaled up ($\times$ 2) to intersect with $\boldsymbol{H}$ sets at 0.17, 0.33, 0.50, 0.67, 0.83. A single concrete pitch is assumed to be interpreted abstractly as harmonic sets (as outlined in Section 3). As such, the A tone on the upbeat before bar 1 supports the root of an *A major* chord set. The A tone as an *A major* chord set extension, and the *crotchet* set extension (upbeat before bar 1) are shown in Fig. 22 (a)–(b). The intersection of these extensions is calculated using the min operator (Eq. (7)), which results in an *A major–crotchet R* time-point relation, shown in matrix form in Fig. 23 (a). The process of performing extensions and intersection to construct the time-point relation *D major–minim S* on bar 1^1 is not shown here, involving the same procedure as that for constructing the first time-point relation, again using Eq. (7). The *D major–minim S* time-point relation is presented in matrix form in Fig. 23 (b).

$$
\begin{array}{l}
\text{[D,F♯,A]} \\ \\ \text{[F♯,A,C♯]} \\ \\ \text{[A,C♯,E]} \\ \\ \text{[C♯,E,G♯]}
\end{array}
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 \\
0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 \\
0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33
\end{pmatrix}
$$

𝅗𝅥 ♩ ♪

(a)

$$
\begin{array}{l}
\text{[D,F♯,A]} \\ \\ \text{[F♯,A,C♯]} \\ \\ \text{[A,C♯,E]} \\ \\ \text{[C♯,E,G♯]}
\end{array}
\begin{pmatrix}
0.33 & 0.50 & 0.67 & 0.83 & 1 & 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\
0.33 & 0.50 & 0.67 & 0.83 & 1 & 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\
0.33 & 0.50 & 0.67 & 0.83 & 1 & 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\
0.33 & 0.50 & 0.67 & 0.83 & 1 & 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\
0.33 & 0.50 & 0.67 & 0.83 & 1 & 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\
0.33 & 0.50 & 0.67 & 0.83 & 1 & 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\
0.33 & 0.50 & 0.67 & 0.83 & 1 & 0.83 & 0.67 & 0.50 & 0.33 & 0.17
\end{pmatrix}
$$

𝅗𝅥 ♩ ♪

(b)

Figure 22: Cylindrical extensions of (a) *A major* and
(b) *crotchet*, “Trout” Quintet (both on upbeat before bar 1).

$$
\begin{array}{l}
\text{[D,F♯,A]} \\ \\ \text{[F♯,A,C♯]} \\ \\ \text{[A,C♯,E]} \\ \\ \text{[C♯,E,G♯]}
\end{array}
\begin{pmatrix}
\mathbf{0} & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.17 \\
0.33 & 0.50 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.50 & 0.33 & 0.17 \\
\mathbf{0.33} & 0.50 & 0.67 & 0.83 & \mathbf{1} & 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\
0.33 & 0.50 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.50 & 0.33 & 0.17 \\
0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.17
\end{pmatrix}
$$

𝅗𝅥 ♩ ♪

(a)

$$
\begin{array}{l}
\text{[D,F♯,A]} \\ \\ \text{[F♯,A,C♯]} \\ \\ \text{[A,C♯,E]} \\ \\ \text{[C♯,E,G♯]}
\end{array}
\begin{pmatrix}
\mathbf{1} & 0.83 & 0.67 & 0.50 & \mathbf{0.33} & 0.17 & 0 & 0 & 0 & 0 \\
0.67 & 0.67 & 0.67 & 0.50 & 0.33 & 0.17 & 0 & 0 & 0 & 0 \\
0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.17 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0} & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

𝅗𝅥 ♩ ♪

(b)

Figure 23: Fuzzy time-point relations for (a) *A major–crotchet R* (upbeat
before bar 1) and (b) *D major–minim S* (bar 1^1), “Trout” Quintet.

To determine the (through-time) R_R, it is then necessary to aggregate time-point relations R and S using the max operator (Eq. (8)). This results in a *D major–minim—A major–crotchet* R_R, shown in Fig. 24 (a), representing both the abstract and concrete R_R (in bold). As noted, the concrete points within the abstract subset relations are the fired time-point relations on the occasion of sense data. The general categories of the matrices represent implicit knowledge of R_R that is presumably innately fixed into the neural architecture of the music faculty. This permits recognition and categorisation of relations when concrete elements in the biological neural network are stimulated by sensory data from the external world. Fig. 24 (b) presents the isolated concrete neurally fired R_R, also satisfying the properties w-reflexivity, w-symmetry and anti-transitivity. A tonal-metrical hierarchy is inferred from the R_R, where the subset relations form each level of the hierarchy: the *D major–minim S* is the higher level, and the *A major–crotchet*

R is the lower level. It should be observed that R_R between non-adjacent time-points are important also for concept-building, and fundamental in the determination of complex hierarchies. Furthermore, it is sometimes necessary that adjacent and non-adjacent sets are *not* connected by R_R, since overabundant R_R would mean hierarchical levels would be set up between a multitude of time-points, which would be perceptually confounding. To this end, logical implication relations (R_M and R_I) enable continuation of fuzzy set structure, extending either the same harmonic set in $\boldsymbol{H}$ or the same grouping level set in $\boldsymbol{G}$ through time, or extending both (Section 5.4).

$$\begin{array}{r} [D,F\sharp,A] \\ \\ [F\sharp,A,C\sharp] \\ \\ [A,C\sharp,E] \\ \\ [C\sharp,E,G\sharp] \end{array} \begin{pmatrix} \mathbf{1} & 0.83 & 0.67 & 0.50 & \mathbf{0.33} & 0.17 & 0 & 0 & 0 & 0 \\ 0.67 & 0.67 & 0.67 & 0.50 & 0.33 & 0.17 & 0 & 0 & 0 & 0 \\ 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.17 \\ 0.33 & 0.50 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.50 & 0.33 & 0.17 \\ \mathbf{0.33} & 0.50 & 0.67 & 0.83 & \mathbf{1} & 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\ 0.33 & 0.50 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.50 & 0.33 & 0.17 \\ 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.17 \end{pmatrix}$$

𝅗𝅥 ♩ ♪

(a)

$$\begin{array}{r} [D,F\sharp,A] \\ [A,C\sharp,E] \end{array} \begin{pmatrix} 1 & .33 \\ .33 & 1 \end{pmatrix}$$

𝅗𝅥 ♩

(b)

Figure 24: (a) Abstract and concrete through-time R_R ($R \cup S$) and (b) the isolated concrete neurally-fired R_R, "Trout" Quintet (Ex. 3).

The perception of low-level R_R is required for individuating low-level voice-leading concepts, such as non-chord tones in serial pitch streams, including passing tones. An unaccented passing tone can be represented as an element within a R_R in a single pitch stream context, where it is a lower-level subset of a low-level R_R (e.g., Ex. 3, bar $3^{4.5)}$. However, in Ex. 1, the non-chord tone E♭ in bar 1^1, termed an accented non-chord tone or appoggiatura in traditional music theory, is introduced at a position of the overall hierarchical structure where a more general R_R is perceived across the whole texture through time, connecting *B♭ major* in $\boldsymbol{H}$ with the *semibreve* level in $\boldsymbol{G}$. If the single E♭ tone in bar 1^1 was to be interpreted abstractly as a harmonic entity it would not provide root support for the more general *B♭ major* chord at this point. Accordingly, the E♭ tone is set apart from the main relation and perceived separately as an accented non-harmonic tone within a different relational structure. Broadly, accented passing tones and appoggiaturas are cases where non-root-supporting tones are introduced that are contradictory to, or form paradoxes with, a wider relational hierarchy, creating conflict or ambiguity against the wider stable conceptual structure. As such, accented passing tones are more commonly found in texturally complex music, not single-line melodies, because they are set up to conflict with more general, higher-level relations (R_R, R_M, or R_I) which provide a background for the smaller paradox (/contradiction), maintaining coherent overall conceptual structure. In more complex situations, where there is consistent conflict through sustained paradoxical relations through time, a novel relational hierarchy is introduced to permit wider comprehensibility (Section 5.5).

In the iterative compositional construction of concepts, relations R_R, R_M and R_I are employed at increasingly higher levels of harmonic–grouping hierarchies until the limit of the perceptual window is reached (to be determined empirically). The highest levels of the product space are the upper bounds of the tonal–metrical hierarchy. A high-level R_R can be seen in Ex. 3, where the *D major–longa* time-point relation in bar 1^1 is aggregated with the *A major–breve* time-point relation in bar 3^1 using the max operator (Eq. (8)). From this R_R, when aggregated with other relations at lower levels, the tonal–metrical structure is extrapolated. Fig. 25 shows the concrete neurally-fired matrix for the *D major–longa—A major–breve* R_R, which satisfies the properties w-reflexivity, w-symmetricality and antitransitivity. In Ex. 3, then, low-level concepts (chords, groupings), mid-level concepts (e.g., chord progressions, harmonic rhythm, passing tones), and high-level concepts (e.g., tonality, metrical structure) are individuated using relations, and assembled according to the principle of compositionality. Low-level complex concepts (chords, groupings) are fixed by low-level relations, and these, in turn, form the building-blocks of mid-level and high-level concepts.

$$\begin{array}{r}[D,F\sharp,A]\\ [A,C\sharp,E]\end{array}\begin{pmatrix}1 & .33\\ .33 & 1\end{pmatrix}$$

Figure 25: Concrete fuzzy R_R, "Trout" Quintet, bb. 1–4.

If comprehension of relations is a perceptual universal, R_R, R_M and R_I should be used to fix concepts across music cultures. The use of R_R can be observed in the opening of a traditional Sundanese gamelan piece, *Lulunga* (Ex. 4). This is an original transcription of a *kecapi suling* setting, performed by Megasari (2020). A *kecapi* is a bamboo flute and a *suling* a type of zither. The Sundanese *pélog* scale, shown in Fig. 26, is not exact in pitch owing to differences with equal temperament tuning, but is approximated using the Phrygian scale on B♭ (Fig. 26). While the modal harmonic system is not fully consistent with the Sundanese harmonic system, it is broadly commensurate: both are constrained by root support, involving scales that incorporate intervals P1/P8, P5 and m3, and they use a similar subset of non-root-supporting intervals. Indeed, most scales of the world are similar to the extent that they incorporate P1/P8, P5 and M3/m3 root-supporting intervals (Gill & Purves, 2009). It seems to be a point of fact that any variation between cultural scales and harmonic systems usually takes place with respect to non-root-supporting interval classes. This tendency in itself is evidence that the harmonic forms of musical cultures are generally constrained by perception based on the criterion of root support.

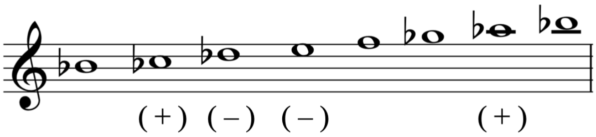
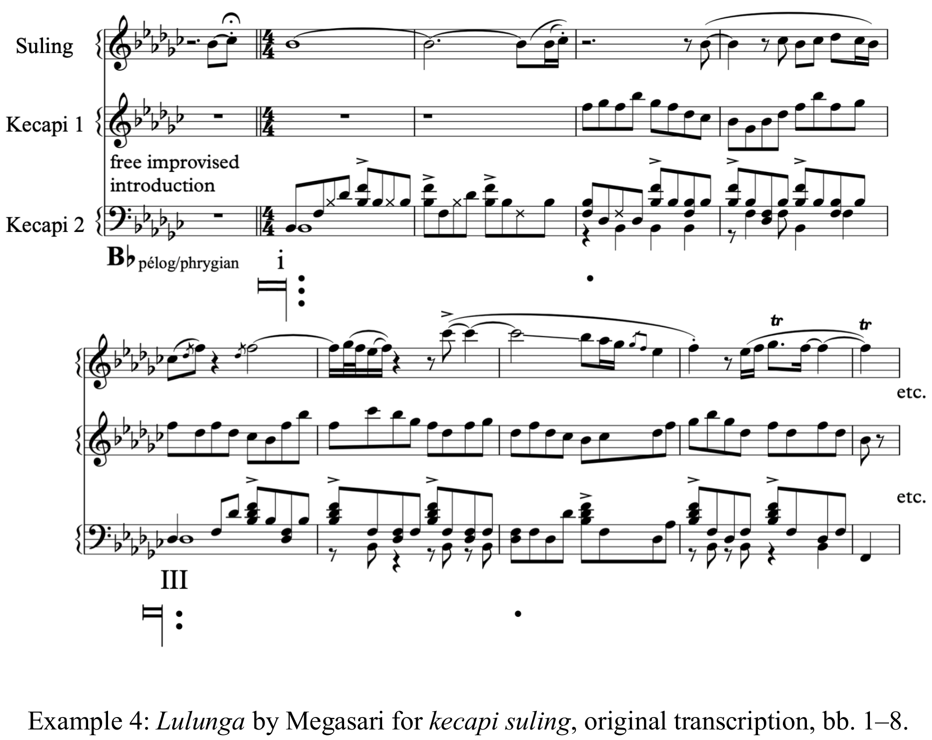


Figure 26: *Pélog* scale on B♭, with cent differences (+ or −) to pitches of equal temperament.



Example 4: *Lulunga* by Megasari for *kecapi suling*, original transcription, bb. 1–8.

At this stage, we do not need to rehearse the process of individuating time-point relations, shown in Eq. (6) and Eq. (7), we can simply posit a *B♭ minor–maxima* time-point relation R at bar 1[1] (Fig. 27 (a)) and a *D♭ major–longa* time-point relation S at bar 5[1] (Fig. 27 (b)). The concrete time-point relations (in bold) are embedded within the abstract time-slice category relations R and S. These are then aggregated using the max operator (Eq. (8)), resulting in an overall *B♭ minor–maxima—D♭ major–longa* R_R. The general category R_R is shown in Fig. 28 (a), with the concrete R_R embedded in bold. The concrete R_R is shown isolated in Fig. 28 (b). The harmonic relationship between the harmonic sets, involving connections by thirds, is traditionally described as Riemannian or neo-Riemannian (Cohn, 1996), or a progression in "thirds space" (Lerdahl, 2001). A disadvantage of these approaches is that they are specialist; they do not generalise between any and all forms of harmonic movement – such as movement by seconds or fifths, for instance. By contrast, ***H*** permits fluid and coherent integration of all types of movement between harmonic sets, and allows coherent integration with ***G***.

$$\begin{array}{l} [B\flat, D\flat, F] \\ \\ [D\flat, F, A\flat] \\ \\ \end{array} \begin{pmatrix} \mathbf{1} & .67 & \mathbf{.33} & 0 \\ .67 & .67 & .33 & 0 \\ \mathbf{.33} & .33 & \mathbf{.33} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(a)

$$\begin{array}{l} [B\flat, D\flat, F] \\ \\ [D\flat, F, A\flat] \\ \\ \end{array} \begin{pmatrix} \mathbf{.33} & .33 & \mathbf{.33} & .33 \\ .33 & .67 & .67 & .67 \\ \mathbf{.33} & .67 & \mathbf{1} & .67 \\ .33 & .67 & .67 & .67 \end{pmatrix}$$

(b)

Figure 27: (a) Abstract and concrete B♭ minor–maxima time-point relation R and (b) abstract and concrete D♭ major–longa time-point relation S (*Lulunga*, bb. 1–8).

$$\begin{array}{l} [B\flat, D\flat, F] \\ \\ [D\flat, F, A\flat] \\ \\ \end{array} \begin{pmatrix} \mathbf{1} & .67 & \mathbf{.33} & .33 \\ .67 & .67 & .67 & .67 \\ \mathbf{.33} & .67 & \mathbf{1} & .67 \\ .33 & .67 & .67 & .67 \end{pmatrix}$$

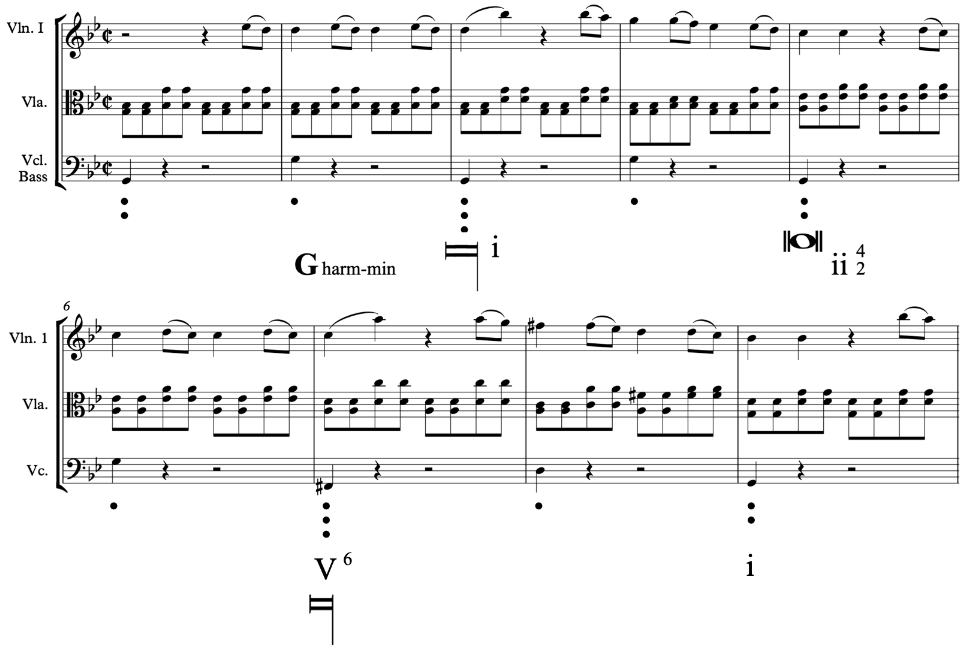
(a)

$$\begin{array}{l} [B\flat, D\flat, F] \\ [D\flat, F, A\flat] \end{array} \begin{pmatrix} 1 & .33 \\ .33 & 1 \end{pmatrix}$$

(b)

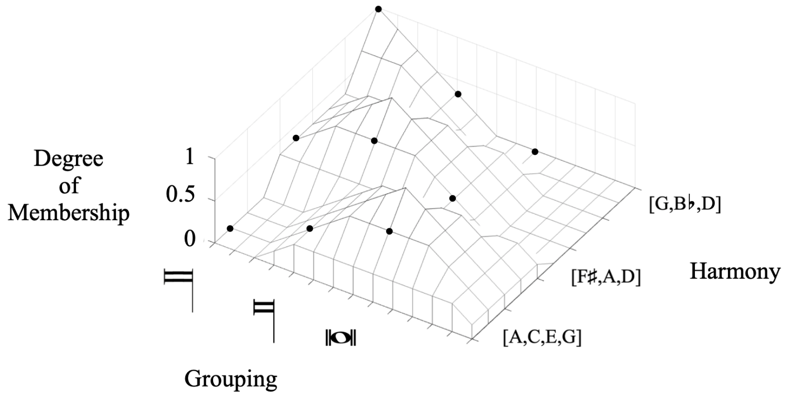
Figure 28: (a) Abstract and concrete R_R (union of R and S) and (b) isolated concrete R_R (*Lulunga*, bb. 1–8).

The opening section of Mozart's Symphony No. 40, i, in Ex. 5 (bb. 1–8) is noteworthy in that it establishes a binary R_R between three time-point subset relations, *G minor–maxima*, *D major–longa* and *A diminished 7th–breve*. When aggregated using the max operator, these form a *G minor–maxima—D major–longa—A diminished 7th–breve* R_R. The concrete R_R (in bold) for this passage is shown embedded within the general fuzzy categories in the matrix in Fig. 29 (a). The concrete R_R within the general categories is again presented in Fig. 29 (b), but in graph form. Note that this example highlights the weakened form of reflexivity, w-reflexivity. The w-reflexivity, symmetry, and the individual concrete subset relations of the tonal–metrical hierarchy of the R_R, are emphasised in a sagittal diagram (Fig. 30 (a)). The concrete R_R is isolated in a fuzzy matrix (Fig. 30 (b)) that highlights all properties: w-reflexivity, symmetry and antitransitivity. It can be seen from these representations that tonality and metrical structure are interdefined by the tonal–metrical hierarchy. The tonal–metrical hierarchy is represented across the whole R_R, synthesising (generalised) root support and (generalised) time regularity. In this example, the highest-level subset is the *G minor–maxima* relation, the next lower-level subset is the *D major–longa* relation, and the lowest-level subset is the *A diminished 7th–breve* relation.

These may be concretised by projection (proj) using Eq. (9), such as reducing the highest levels of tonality and metrical structure to the G minor set and maxima level set, respectively. A further possible stage in the cognition of R_R is that a complex relation may be defuzzified to enable crisp and efficient classification of sets and relations, for simplicity and efficiency of processing. As discussed, crisp concepts reduce memory load, enabling more efficient top-down processing, and support coherent connections between other cognitive systems, such as consciousness, and belief and memory systems. Defuzzification allows competing or contradictory fuzzy truth values or truth functions to be synthesised, since fuzzy logic invalidates the "laws of thought", i.e., law of identity, law of the excluded middle and law of non-contradiction (Fodor, 1975, 1998, 2008), hindering some forms of cognitive processing. Standard methods of defuzzification include alpha cuts, mean–max membership, the centroid method and weighted average (Klir & Yuan, 1995). For instance, a defuzzification of Ex. 5 can be made using an alpha cut at $\alpha = .67$, which simplifies the fuzzy R_R to a crisp bivalent R_R (Fig. 30 (c)). If required, crisp representation can be made at several alpha cuts, such as $\alpha = .33$ and $\alpha = 0$ (not shown). Such concerns relating to defuzzification of relations require further investigation in future work.

Example 5: Harmony sets (Roman numerals) and
grouping sets (maxima and longa) in
Symphony No. 40, i, W.A. Mozart, bb. 1–9 (1788).

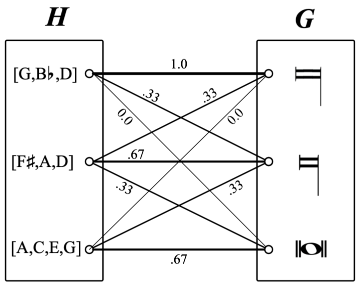
$$
\begin{array}{l}
[G,B\flat,D] \\
\\
[B\flat,D,F] \\
\\
[D,F\sharp,A] \\
[F\sharp,A,D] \\
[F\sharp,A,C] \\
\\
[A,C,E] \\
[A,C,E,G] \\
\\
\end{array}
\begin{pmatrix}
\mathbf{1} & 0.83 & 0.67 & 0.50 & \mathbf{0.33} & 0.17 & 0 & 0 & \mathbf{0} \\
0.67 & 0.67 & 0.67 & 0.50 & 0.33 & 0.17 & 0 & 0 & 0 \\
0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 \\
0.33 & 0.50 & 0.67 & 0.67 & 0.67 & 0.67 & 0.67 & 0.50 & 0.33 \\
0.33 & 0.50 & 0.67 & 0.83 & 1 & 0.83 & 0.67 & 0.50 & 0.33 \\
\mathbf{0.33} & 0.50 & 0.67 & 0.67 & \mathbf{0.67} & 0.67 & 0.67 & 0.50 & \mathbf{0.33} \\
0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33 \\
0 & 0 & 0 & 0.17 & 0.33 & 0.50 & 0.67 & 0.67 & 0.67 \\
0 & 0 & 0 & 0.17 & 0.33 & 0.50 & 0.67 & 0.83 & 1 \\
\mathbf{0} & 0 & 0 & 0.17 & \mathbf{0.33} & 0.50 & 0.67 & 0.67 & \mathbf{0.67} \\
0 & 0 & 0 & 0.17 & 0.33 & 0.33 & 0.33 & 0.33 & 0.33
\end{pmatrix}
$$

(a)



(b)

Figure 29: (a) Fuzzy abstract and concrete *G minor–maxima—D major–longa—A diminished 7th–breve* R_R, on a matrix within general categories of each domain and (b) graph showing fuzzy concrete R_R within the abstract categories of each domain.



$$
\begin{array}{l}
[G,B\flat,D] \\
[F\sharp,A,D] \\
[A,C,E,G]
\end{array}
\begin{pmatrix}
1 & .33 & 0 \\
.33 & .67 & .33 \\
0 & .33 & .67
\end{pmatrix}
\qquad
\begin{array}{l}
[G,B\flat,D] \\
[F\sharp,A, D] \\
[A,C,E,G]
\end{array}
\begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}
$$

(a) (b) (c)

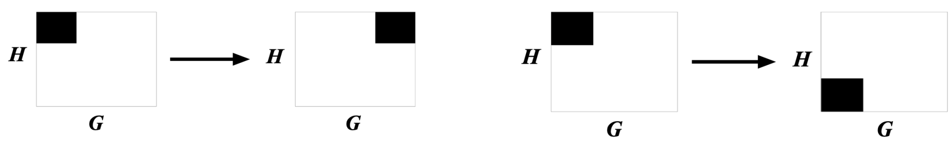
Figure 30: (a) concrete fuzzy R_R in sagittal diagram,
(b) concrete fuzzy R_R in matrix diagram and
(c) concrete crisp alpha cut (α = .67) of R_R.

5.4 Fuzzy Implication Relations

Fuzzy relational implication concerns the boundedness or overlap through time between sets of one domain (R_M) or both domains (R_I) in the $\boldsymbol{H} \times \boldsymbol{G}$ product space. These forms of logical implication enable elements, or relations between elements of the domains, to be perceptually connected through time for the construction of truth-functional strings in $\boldsymbol{H}$, $\boldsymbol{G}$, or both domains combined. This subsection focuses on the individuation and iterative assembly of R_M and R_I.

Fuzzy implication relations (R_M and R_I) only obliquely correspond with notions of implication or prolongation espoused in music theory and computational musicology (as discussed in Section 1.1) (cf. Schenker, 1935; Meyer, 1956; Lerdahl & Jackendoff, 1983; Narmour, 1990; Lerdahl, 2001; Hamanaka et al., 2006; Lerdahl & Krumhansl, 2007; Marsden, 2010; Marsden et al., 2018). Generative theory categorises harmony based on prolongational connections, even when harmonic connections may only be tenuous or vague between entities through time. In FRMP, by contrast, implication is defined strictly as overlap or continuation of conceptual structure within a domain or in both domains of the product space through time. Where there is no fuzzy overlap in a domain, such as when novel sets are introduced in $\boldsymbol{H}$ or $\boldsymbol{G}$, this state of affairs is interpreted as fuzzy negation (/complement) of a prior term, using Eq. (4), and results in concept change. When negation occurs in both domains, new levels of conceptual hierarchy are introduced by leveraging R_R; or, if the subset is irreconcilable within the relational hierarchy, a novel tonal–metrical hierarchy is introduced.

To determine fuzzy implication in a single domain (R_M), a standard t-norm operator is required; presently, the min operator is elected (Eq. (5)). R_M is central to the classification of many concepts, such as where a string of tones through time can implicate a unified chord by virtue of support for a single harmonic root. Classification of harmony through time requires implication between $\boldsymbol{H}$ elements based on root support. *Mutatis mutandis*, implication between $\boldsymbol{G}$ elements based on time regularity is necessary to perceive and classify repeated regular grouping sets through time. The use of R_M in the $\boldsymbol{H}$ domain to individuate a chord concept can be seen in the opening passage of Schubert's "Trout" Quintet (Ex. 3), which involves harmonic implication between beats 1, 3 and 4 of every bar. In bar 1, for instance, all pitch elements of the melody and accompaniment on beats 1, 3 and 4 strongly implicate the same root pitch D, thus collectively individuating the *D major* chord set. The heatmaps in Fig. 31 (a)–(b) show schematic diagrams of single-domain relational implication R_M for $\boldsymbol{H}$ (a) and $\boldsymbol{G}$ (b), with varying conceptual structure in the co-domain. Thus, for implication within a domain, harmonic or grouping sets must remain *the same* or *similar* at a particular hierarchical level during the unfolding of events. This is experienced in perception as the continuation of a chord set or the continuation of a grouping set. The fuzzy implication calculus is therefore an incisive means of measuring perceptual similarity in music (cf. Marsden 2012). A R_M is also important for harmonic change, in connection with the converse of implication, negation, to define a novel instantiation event, which is a constraint on time regularity in domain $\boldsymbol{G}$ (constraint *a*). Negation enables further explication of the individuation of passing tones, which can be defined as harmonically non-implicated (non-chord/negated) tones that move by step between two harmonically-implicated (chord) tones. The model of passing tones as non-implicated (/negated) tones is coherent with that which explains passing tones as a lower-level subset of R_R, outlined above.



(a) (b)

Figure 31: Schematic heatmaps of R_M in (a) $\boldsymbol{H}$ and (b) $\boldsymbol{G}$.

Relational implication in both domains simultaneously (R_I) involves overlap of subsets in both domains in the product space through time, permitting continuation of both sets in the tonal–metrical hierarchy. Eq. (15) expresses R_I, which involves both domains, where the intersection is the relational subset X and Y in the $\boldsymbol{H} \times \boldsymbol{G}$ product space through time ($X \rightarrow Y$), leveraging the Mamdami min operator (Mamdami and Assilian, 1975).

$$R_I = \mu_{\underset{\sim}{X}}(h, g) \rightarrow \mu_{\underset{\sim}{Y}}(h, g) = \min\left(\mu\underset{\sim}{X}(h, g), \mu\underset{\sim}{Y}(h, g)\right) \tag{15}$$

Schubert's "Trout" Quintet (Ex. 3) features R_I between adjacent and non-adjacent time-point relations. For example, the *semibreve–D major* time point relation in bar 1^1 is implicated in a virtual repetition of that time-point relation in bar 2^1, duplicating the tonal–metrical hierarchy at the latter adjacent time-slice. As such, the *semibreve–D major* time-point relation in bar 1^1 occupies a similar space in the product space as the time-point relation in bar 2^1, determined by Eq. (15). There are also low-level, non-adjacent R_I in Ex. 3, such as R_I between notated beats 2 and 4 of every bar. For instance, the subset *D major–crotchet* in bar 1^2 is implicated in bar 1^4, duplicating the tonal–metrical hierarchy at this point. Fig. 32 shows a schematic heatmap for R_I, where there is complete duplication (overlap) in the product space. Note that R_I may likewise be negated, which can be calculated by substitution of the MC terms of R_I into Eq. (4).

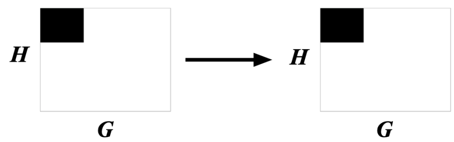
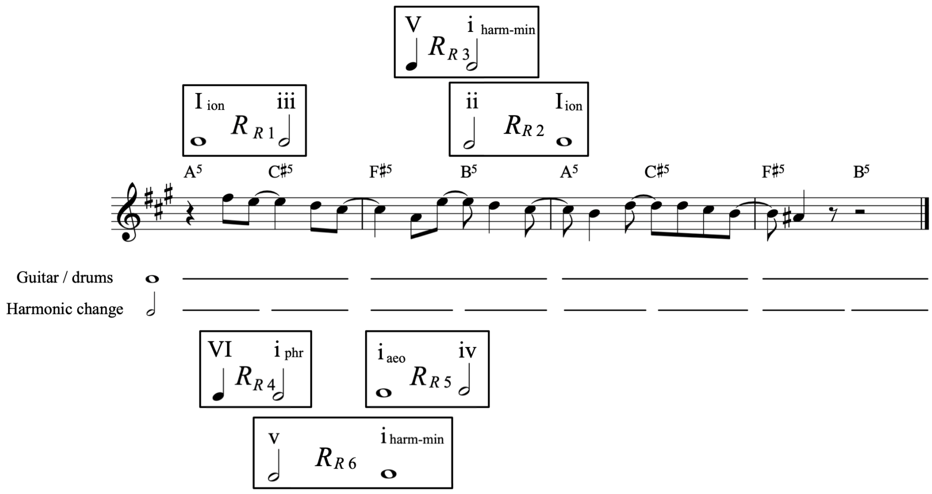


Figure 32: Schematic heatmap showing R_I.

5.5 Plural Relational Hierarchies

With the introduction of a novel set or relation, a novel relational hierarchy may be initiated to avoid paradox or contradiction with a prior relational hierarchy. This state of affairs is termed plural relational hierarchies. Plural relational hierarchies are constructed in various types of music, including the multiple independent melodic streams of polyphony or counterpoint, broadly conceived, or even within a single melodic stream of a song. Indeed, locked inside a melodic stream are myriad multiparametric concepts, from basic to complex, involving harmony, grouping, harmonic rhythm, etc., which may combine in various ways to form separate, or even contradictory, relational hierarchies. Each hierarchy is presented on a separate $\boldsymbol{H} \times \boldsymbol{G}$ product space and processed independently in the perceptual music faculty. However, the integration of hierarchies, which may take place in cognition, is opaque to formal analysis at present (cf. Chomsky, 2009). This section aims to sketch a tentative story on the broad picture of plural hierarchical representation.

The song "Drain You" (1991) by the rock band Nirvana (Ex. 6), has been suggested to have plural harmonic–grouping (/tonal–metrical) hierarchical structures (Rawbone, 2021). The song contains six resemblance relations, $R_{R\,1\text{–}6}$, which are differentially coherent. The matrices for $R_{R\,1\text{–}6}$ are not illustrated here, but we can simply posit R_R based on the hierarchical relational theory above, configuring the domains to the appropriate harmonic roots and grouping levels, scaled to permit coherent interaction on the product spaces. Let us firstly examine the fuzzy sets formed in $\boldsymbol{G}$. Instantiation articulates crotchet-level sets [.25] at notated crotchets by all instruments, although in particular, in the melody stream (constraint *a*). Fuzzy minim level grouping sets [.5] are individuated by harmonic instantiation, through harmonic change at the notated bar and half-bar onsets (constraint *a*). Harmonic instantiation in $\boldsymbol{G}$ occurs with reference to information from $\boldsymbol{H}$, where harmony is negated at the onset of every harmony (Eq. (4)) and novel harmony established through time by implication (R_M) (Eq. (5)). Fuzzy semibreve-level grouping sets [1] at the notated bar are also individuated by instantiation through harmonic negation and subsequent implication (constraint *a*), by relative textural density in guitar and drums at the onset of notated bar time-points (constraint *b*) and through parallel rhythms corresponding at time-points marked by the notated bar (constraint *d*). Ex. 6 is noteworthy in that it involves intersectional aggregation in $\boldsymbol{G}$ of semibreve-level and minim-level sets at notated half-bar time-points, shown in the partition structure in Fig. 33. That is, the fuzzy semibreve-level grouping sets [1] intersect with fuzzy minim-level sets [.5] at notated bar and half-bar time-points to form a combined intersect. This semibreve–minim intersectional set [1,.5] presumably has a membership value somewhere just under the λ_1 level, owing to the strong component of the semibreve level subset, i.e., [1]. Significantly, no level sets higher than the semibreve level are articulated in $\boldsymbol{G}$, and as such there are no R_R at level sets higher than the semibreve level, as will be shown in the ensuing analysis.



Example 6: $R_{R1\text{–}6}$ in “Drain You”, Nirvana (1991).

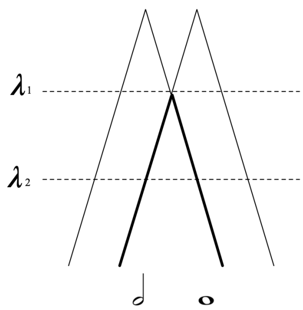


Figure 33: Intersection of semibreve and minim
grouping sets at the notated half-bar time-points in “Drain You”.

The relational interaction between domains generates conflicting plural hierarchies. The $R_{R1\text{–}6}$ in the first harmonic loop (bars 1–2) are identical to the second 2-bar loop (bars 3–4) in Ex. 6, but for brevity the relation notation is not repeated. As noted, $R_{R1\text{–}6}$ are differentially coherent with each other, and the conflict between the relations means higher-level tonal–metrical structure is not intimated. The relations R_{R1} and R_{R2} are both *A major–semibreve—C♯ minor–minim* R_R, which can be concretised using projection (Eq. (9)) to A major/Ionian key at the semibreve level of tonal–metrical structure, on the onset of notated bars 1 and 3. Also, R_{R5} and R_{R6} are both *F♯ minor/major–semibreve—B minor–minim* R_R, which can be concretised to F♯ Aeolian/harmonic major at the semibreve level of metrical structure, on the onset of notated bars 2 and 4. However, R_{R3} is a *B minor–minim—F♯ minor/major–crotchet* R_R, suggesting B Aeolian/harmonic minor at the minim level of metrical structure, on the notated half-bar time-points of bars 2 and 4. Also, R_{R4} is a *C♯ minor–minim—A major–crotchet* R_R, resulting in C♯ minor at the minim level of metrical structure, at the notated half-bar time-points of bars 1 and 3. Thus, R_{R3} and R_{R4} cue a harmony (/tonality) and weak grouping (/metrical) hierarchy at notated minim time-points (shown in Fig. 33). These harmony–grouping relations and their suggested tonal–metrical hierarchies conflict with each other at the notated bar level and higher. Accordingly, they do not permit perception of a coherent monistic higher-level tonal–metrical hierarchy.

While plural relational hierarchies are a paradigm feature of Renaissance and Baroque counterpoint, they are also fundamental to many other types of Western and non-Western musics. For example, pluralistic hierarchies underpin Sundanese gamelan textures (e.g., Ex. 4) and Classical-period structures (e.g., Ex. 5), among many other styles of many cultures and historical periods throughout the world. Ex. 4 (*Lulunga*) involves a plurality of tonal–metrical hierarchies constructed from the three individual instrumental parts. However, the above analysis of Ex. 4 reflects only the generalised overall hierarchy, largely informed by the accompanimental zither (*suling*). A deeper examination would reveal the plural hierarchies that exist within this generalised hierarchy. The analysis of Ex. 5, the opening of Mozart's Symphony No. 40, i (1788), concerns an overall tonal–metrical hierarchy likewise informed by the accompanimental instrumental parts; however, there is also a superimposed melodic tonal–metrical relational hierarchy presented by the violin I part, which is out of phase with the main orchestral accompaniment (Bernstein, 1976; Lerdahl and Jackendoff, 1983). Such phase differences have been characterised as "extended anacruses" (Rothstein 1989; McKee 2004), but actually concern a more fundamental perceptual capacity for individuating independent harmonic–grouping hierarchies. Plural relational hierarchies require teasing apart in perception, because they are an essential property of musical structure caused by the top-down agency of composers. A further case should be mentioned, which involves situations where a melody stream is accompanied by an additional stream or texture that merely makes explicit conceptual structure already presented in the main melodic stream. Such textures, found in popular music settings, hymn tune keyboard parts Western and non-Western folk music arrangements, for example, do not generate bone fide pluralistic relational hierarchies but converge into a single overall harmonic–grouping (tonal–metrical) hierarchy.

The explanation of plural relational hierarchies provided counters the conviction that only *one single metre* or *one overall tonality* is held in perception (/cognition) during real-time listening (cf. Krumhansl, 1990; Conklin & Witten, 1995; Lerdahl, 2001; London, 2004; Hamanaka et al., 2006; Temperley, 2006; Pearce & Wiggins, 2012; Todd, 2015; Pearce, 2016; Marsden et al., 2018). Plural fuzzy relational hierarchies are common in many musical cultures, and so the capacity to generate them must be innate to perception. They may also be a method of admitting structural ambiguity into music, which seems to be a factor of artistic creativity (Bernstein, 1976).

6 Conclusions: Fuzzy Relational Music Perception in Context

FRMP is a formal fuzzy mathematical model showing how low-level concepts (e.g., harmony and grouping), mid-level concepts (e.g., harmonic rhythm) and high-level concepts (e.g., tonal and metrical structure) are constructed combinatorially and compositionally in perception by relations between $\boldsymbol{H}$ and $\boldsymbol{G}$. Domains $\boldsymbol{H}$ and $\boldsymbol{G}$ comprise distinct partition structures based on generalised root support and generalised time regularity, respectively, which are connected by fuzzy relations of resemblance (R_R) and implication (R_M and R_I). These relations are decoded by perception for the individuation and assembly of conceptual structure. R_R connect both domains through time, comprising the properties w-reflexivity, w-symmetry and antitransitivity, and R_M and R_I connect one or both domains through time, respectively, involving the continuation or overlap of concepts. Using these fuzzy relations, it has been theorised and demonstrated that perception generates coherent formulations of the following universal concepts and structural types: harmony, grouping, tonal structure, metrical structure, chord inversion, harmonic progression, added chord tones, modulation, chromaticism, counterpoint, accented and unaccented passing tones, harmonic change and harmonic rhythm, among others.

Concepts determined by R_R, R_M and R_I are arranged hierarchically through time, resulting in theoretically unbounded conceptual hierarchies. The music faculty, which is the perceptual module that individuates these relations, thus concerns a set-theoretical and truth-functional musical thought language (Chomsky, 1957, 1966; Fodor, 2008; Rawbone, 2021). FRMP amounts to an innate, graded and intrinsic form of musical semantics, mirroring formal semantics rather than generative syntax, but incorporating fuzzy propositional logic rather than the predicate logics and higher-order logics of formal semantics (cf. Portner & Partee, 2002). The concepts and conceptual hierarchies of FRMP may be interpreted by more general modules at higher levels of cognition but such processes are not understood at present. Cognitive processing has not been given focus, although it may be assumed that modules such as consciousness, belief systems and memory systems are essential for interpreting meaning in emergent perceptual edifices.

FRMP offers a framework for determining MC relations and concept construction that is more compact, direct and parsimonious than generative theories, avoiding meta-conceptual notions such as time-span reduction and prolongational reduction, which can result in systemic fragmentation and incoherence. FRMP is also a marked alternative to connectionist, statistical–associative and inductive frameworks, because empirically-guided processes often involve entrenched constellations of elements that do not account for the graded and combinatorially unbounded hierarchies of musical structure. The concepts and conceptual hierarchies of music have been shown to be generated as a result of the requirement to form fuzzy relational interconnections between domains $\boldsymbol{H}$ and $\boldsymbol{G}$. Without a principled basis for interaction between the domains, comprehension of a musical thought language would be impossible. This view obliquely accords with Chomsky's (1957) classic argument for generativity of language, which determines the limitations of associative–statistical methods in parsing unbounded combinatorial structures in natural languages (cf. Meyer, 1956; Gjerdingen, 1988; Conklin & Witten, 1995; Temperley, 2006; Gjerdingen, 2007; Bharucha, 2009; Pearce & Wiggins, 2012; Hansen & Pearce, 2014; Dhariwal et al., 2020). If music perceptual comprehension is innate, which is suggested by FRMP, prior learning or exposure to lexicons or data sets may not be necessary for music processing. That is, learning, probability and expectation may not be essential conditions of music perception, although statistical information seems to be incorporated, albeit enigmatically, into higher-level cognitive processing.

The limitations of associative–statistical and connectionist models may be extrapolated from the above analysis of the opening of Mozart's Symphony No. 40, i (1788) (Ex. 5). This passage has been described by schema theorists as a "Meyer" schema (Gjerdingen, 1988, 2007) and also as a "Gavotte" schema (Mirka, 2009), based on the statistical combination of its features. However, it has alternatively be interpreted as an MC arrangement of cognitively-generated universal features, termed a "butterfly" schema (Rawbone & Jan, 2020; see also Lerdahl, 2001). FRMP enables the MC "butterfly" explanation to be more finely formalised. The passage can be explained as a three-element R_R (Fig. 29 and Fig. 30), representing the first three chord sets and their corresponding grouping sets (bars 3–8), with the addition of a R_M between the initial and final tonic chord sets (bar 3 and bar 9). It seems perceptual representation or categorisation of such schemas does not require knowledge of the *cultura galante* (although knowledge of such cultural forms may be an actual property of higher-level cognition), or any prior statistical learning, but uses innate knowledge of fuzzy perceptual relations between $\boldsymbol{H}$ and $\boldsymbol{G}$ that is already fixed in neural architecture. Broadly, the majority of so-called schemas can be individuated a priori using R_R, R_M, or R_I, where relations are constructed across domains to connect root support in $\boldsymbol{H}$ with time regularity in $\boldsymbol{G}$. This claim encompasses stock musical cadences, such as the perfect cadence, which may simply be reduced to a fuzzy essence comprising a R_R between $\boldsymbol{H}$ and $\boldsymbol{G}$, where a chord I follows a chord V in $\boldsymbol{H}$ at two time-points at relatively high levels of $\boldsymbol{G}$. Note that this position is a distinct contrast to the anti-essentialist positions of Gjerdingen (2007) and Rawbone & Jan (2020). The present approach may be considered *graded essentialist*, since mental representations are limited by innate graded essences based on (generalised) root support and (generalised) time regularity in $\boldsymbol{H}$ and $\boldsymbol{G}$, and the graded relational possibilities between them.

A main limitation of FRMP as it currently stands is that it does not provide formal systems for constructing membership values given the partition structures and the domain constraints. Conceiving an algorithm that synthesises the constraints for $\boldsymbol{H}$ and $\boldsymbol{G}$ using the partition structures and outputs suitable membership values is thus an important avenue for future research. Perhaps a further limitation of the model is that it offers only an ancillary formulation of chromaticism and modulation. It is foundational that root-detracting harmony must give rise to novel chords (involving a novel root) or novel $\boldsymbol{H}$ domains (involving a novel tonic), as illustrated in "I Get Around" by the Beach Boys (Ex. 2). Chromaticism and modulation thus do not have their own formal frameworks as such but emerge as properties that are by-products of an absence of root support. This is a markedly different approach to many voice-leading and geometry theories, where chromatic movement and tonal geometries are given concrete representation using systematic frameworks that treat these phenomena on their own terms (e.g., Cohn, 1996; Tymoczko, 2012). The present indirect framing of chromaticism and modulation may be preferable, however, since it is contestable whether non-logical properties should form part of a formal fuzzy-logical framework. Assuming the present set-theoretical and logical semantics, it would be incomprehensible to present chromatic and modulatory phenomena as basic conceptual infrastructure, because such have been shown to be intrinsically non-truth-functional. If the present approach is a true approximation of music perception, comprehension of emergent chromatic elements and modulatory structure must involve reframing within higher-level cognitive systems by top-down general systems, although the emergent phenomena would always be intractable to perception. Ultimately, modulation and chromaticism may be considered open problems of this research program.

Plural relational hierarchies are set up to avoid paradox or contradiction, to parse a term that is non-congruent in a prior hierarchy. The evidence for plural relational hierarchies challenges the view that there are necessarily monistic tonal and metrical hierarchies (cf. Lerdahl & Jackendoff, 1983; Lerdahl, 2001; London, 2004; Deutsch 2013b; Tojo et al., 2018). Some form of synthesis or rationalisation of plural relational hierarchies – presumably by higher-level systems in cognition – may be required to determine the interaction between them. This notion warrants further investigation, since at present, the connection (if any) between relational hierarchies is not understood, although may be presumed to inform meaning in unique artistic music. A goal of future research may be to map out and provide generalisations for the rich hierarchies locked inside serial pitch streams (e.g., melody), where parametric features such as harmony, grouping, tonality and metrical concepts form independent hierarchies within more general hierarchies. The role of more peripheral parametric structures may also be examined, such as contour, chromaticism, dissonance and voice-leading, etc. With respect to the latter properties, the work of Yilmaz & Teletar (2010a, 2010b) on counterpoint rules may provide useful formulations from which to proceed.

Analogical resemblance relations (R_R), which involve w-reflexivity, w-symmetry and antitransitivity across binary fuzzy domains, may not only be of significance for music perception, but could have general application as a broad psychological principle for codifying analogy between disparate entities in perception, which is a hypothesis that requires further investigation. Future research may aim to corroborate FRMP empirically, through observational approaches in neuroscience, participant testing in psychology, and formal implementation in computer science. Toward computational implementation, as discussed, it would be desirable to construct an algorithm that, given an input of musical parameters (e.g., MIDI), generates membership values based on the partition structures and domain constraints. It would also be advantageous to more clearly delineate the characteristics and functions of plural relational hierarchies, as suggested. The perception of plural hierarchies seems to be connected, in an oblique way, to mental complexity and creativity, and is an area where fuzzy models may offer considerably greater insight than the relatively static picture of music cognition espoused by generative and associative–statistical theories.

7 References

Alsinet, T., & Godo, L. (2000). A complete calculus for possibilistic logic programming with fuzzy propositional variables. In C. Boutilier & M. Goldszmidt (Eds.), *Proceedings of the conference on uncertainty in artificial intelligence 2000* (pp. 1–10). Morgan Kaufmann.

Armengol, E., Dellunde, P., & García–Cerdaña, À. (2015). A logical study of local and global graded similarities. *Applied Artificial Intelligence*, *29*, 424–444.

Azad, F., & Shukla, P. K. (2021). A review on the development of fuzzy classifiers with improved interpretability and accuracy parameters. *Journal of Informatics, Electrical and Electronics Engineering*, *2*(2), 1–9.

Bach, E. (1964). *An introduction to transformational grammars*. Holt, Rinehart and Winston.

Balzano, G. J. (1982). The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, mind, and brain* (pp. 321–351). Plenum Press.

Bernstein, L., (1976). *The unanswered question: Six talks at Harvard*. Harvard University Press.

Beg, I., & Ashraf, S. (2010). On Poincaré paradox. *The Journal of Fuzzy Mathematics*, *18*(3), 631–640.

Benjamin, W. E. (1984). A theory of musical meter. *Music Perception*, *1*, 355–413.

Bharucha, J. J. (2009). From frequency to pitch, and from pitch class to musical key: Shared principles of learning and perception. *Connection Science, 21*(2–3), 177–192.

Bregman, A. S. (1990). *Auditory scene analysis*. MIT Press.

Buchler, M. (2001). Relative saturation of intervals and set classes: A new approach to complementation and pc set similarity. *Journal of Music Theory*, *45*(2), 263–343.

Cádiz, R. F. (2020). Creating music with fuzzy logic. *Frontiers in Artificial Intelligence*, *3*(59), 1–20.

Chew, E. (2001). Modeling tonality: Applications to music cognition. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the twenty-third annual conference of the cognitive science society* (pp. 206–212). Lawrence Erlbaum Associates.

Chomsky, N. (1957). *Syntactic structures*. Mouton.

Chomsky, N. (2009). (1966). *Cartesian linguistics*: *A chapter in the history of rationalist thought*. Cambridge University Press.

Chomsky, N. 1995. *The minimalist program*. MIT Press.

Chon, I. (2003). G-fuzzy equivalence relations generated by fuzzy relations. *Korean Journal of Mathematics*, *11*(2), 169–175.

Chon, I. (2017). Weak fuzzy equivalence relations and weak fuzzy congruences. *Korean Journal of Mathematics*, *25*(4), 563–577.

Civit, M., Civit–Masot, J., Cuadrado, F., & Escalona, M. (2022). A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends. *Expert Systems with Applications*, *209*, 1–16.

Cock, D. M., Bodenhofer, U., & Kerre, E. E. (2000). Modelling linguistic expressions using fuzzy relations. In *Proceedings of the 6th international conference on soft computing* (pp. 353–360). Fuzzy Logic Systems Institute.

Cock, D. M., & Kerre, E. E. (2001). Approximate equality is no fuzzy equality. In S. Massanet, S. Montes, D. Ruiz-Aguilera, & M. Gonzáles-Hildago (Eds.), *Proceedings of the 2nd international conference in fuzzy logic and technology* (pp. 369–371). Atlantis Press.

Cock, D. M., & Kerre, E. E. (2003a). On (un)suitable fuzzy relations to model approximate equality. *Fuzzy Sets and Systems*, *133*(2) 137–153.

Cock, D. M., & Kerre, E. E. (2003b). Why fuzzy *T*-equivalence relations do not solve the Poincaré paradox, and related issues. *Fuzzy Sets and Systems*, *133*(2), 181–192.

Cohn, R. (1996). Maximally smooth cycles, hexatonic systems, and the analysis of late-romantic triadic progressions. *Music Analysis*, *15*(1), 9–40.

Cohen, M. A., Grossberg, S., & Wise, L. (1995). A spectral model of pitch perception. *The Journal of the Acoustical Society of America*, *98*, 862–879.

Cone, E. T. (1968). *Musical form and musical performance*. Norton.

Conklin, D., & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, *24*(1) 51–73.

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. University of Chicago Press.

Dhariwal, P., Jun, H., Payne, C., Kim, J. W., Radford, A., & Sutskever, I. (2020). JukeBox: A generative model of music. *ArXiv*. https://www.openai.com/blog/jukebox.

Deliège, F., & Pedersen, T. B. (2007). Fuzzy song sets for music warehouses. In S. Dixon, D. Bainbridge, R. Typke (Eds.), *Proceedings of the 8th international conference on music information retrieval*, *ISMIR* 2007 (pp. 21–26). International Society for Music Information Retrieval.

Deutsch, D. (1998). The psychology of hearing. *Sound & Video Contractor*. https://www.svconline.com/news/psychology-hearing-365745.

Deutsch, D. (2013a). Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music* (3rd ed.) (pp. 183–249). Academic Press.

Deutsch, D. (2013b). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (3rd ed.) (pp. 249–325). Academic Press.

Elsea, P. (1995). *Fuzzy logic and musical decisions*. http://artsites.ucsc.edu/ems/Music/research/FuzzyLogicTutor/FuzzyTut.html.

Fitzgerald, J. A., Geiselbrechtinger, F., & Kechadi, T. (2004). Feature extraction of handwritten symbols using fuzzy logic. In G. I. McCall (Ed.), *Proceedings of the 11th biennial conference of the Canadian society for computational studies of intelligence on advances in artificial intelligence* (pp. 493–498). Springer-Verlag.

Fodor, J. A. (1975). *The language of thought*. Harvard University Press.

Fodor, J. A. (1983). *The modularity of mind*: *An essay on faculty psychology*. MIT Press.

Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford University Press.

Fodor, J. A. (2008). *LOT 2*: *The language of thought revisited*. Oxford University Press.

Forte, A. (1973). *The structure of atonal music*. Yale University Press.

Garriga–Berga, C. (2005). *A new approach to the synthesis of intelligible fuzzy models from input–output data* [Unpublished doctoral thesis]. Universitat Ramon Llul.

Gill, K. Z., & Purves, D. (2009). A biological rationale for musical scales. *PloS ONE*, *4*(12), e8144.

Gjerdingen, R. (1988). *A classic turn of phrase*: *Music and the psychology of convention*. University of Pennsylvania Press.

Gjerdingen, R. (2007). *Music in the galant style*. Oxford University Press.

Golik, P., Harb, B., Misra, A., Riley, M., Rudnick, A., & Weinstein, E. (2012). Mobile music modeling, analysis and recognition. In A. Sugiyama, H. Kiya, & H. Sakai (Eds.), *IEEE international conference on acoustics*, *speech*, *and signal processing, ICASSP* (pp. 2353–2356). IEEE.

Gonzalez-Inostroza, M., de La Cuadra, P., & Cádiz, R. F. (2015). Fuzzy equalization of musical genres. In R. Dudas (Ed.), *41st International computer music conference*: *Looking back, looking forward – proceedings* (pp. 134–137). International Computer Music Association.

Griffiths, T. D., & Frackowiak, R. (1998). Analysis of temporal structure by the human brain. *Nature Neuroscience*, *1*(5), 422–427.

Gupta, C., Jain, A., & Joshi, N. (2018). Fuzzy logic in natural language processing: A closer view. *Procedia Computer Science*, *132*, 1375–1384.

Gupta, K. C., & Gupta, R. K. (1996). Fuzzy equivalence relation redefined. *Fuzzy Sets and Systems*, *79*, 227–233.

Hamanaka, M., Tirata, K., & Tojo, S. (2005). ATTA: Automatic time-span tree analyzer based on extended GTTM. In J. Reiss (Ed.), *ISMIR 2005 international conference on music information retrieval*: *Proceedings* (pp. 358–365). Queen Mary University of London.

Hamanaka, M., Tirata, K., & Tojo, S. (2006). Implementing "A generative theory of tonal music". *Journal of New Music Research*, *35*, 249–277.

Hamanaka, M., Tirata, K., & Tojo, S. (2007). FATTA: Full automatic time-span tree analyzer. In K. K. Jensen & S. Serafin (Eds.), *Proceedings of the international computer music conference* (pp. 153–156). International Computer Music Association.

Hamanaka, M., & Tojo, S. (2009). Interactive GTTM analyzer. In K. Hirata, G. Tzanetakis, & K. Yoshii (Eds.), *Proceedings of the 10th international society for music information retrieval conference* (pp. 291–296). International Society for Music Information Retrieval.

Hansen, N. C., & Pearce, M. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology*, *1*, 1–17.

Harnad, S. (1990). The symbol-grounding problem. *Physica*, *42*, 335–346.

Hasanzadeh, F., Annabestani, M., & Moghimi, S. (2019). Continuous emotion recognition during music listening using EEG signals: A fuzzy parallel cascades model. *ArXiv*. 1910.10489.

He, Y., & He, P. (2019). Design of artificial music intelligence system based on fuzzy perception learning. *Journal of Physics*, *1237*, Article 042031. https://doi:10.1088/17426596/1237/4/042031.

Hüllermeier, E. (2015). Does machine learning need fuzzy logic? *Fuzzy Sets and Systems*, *281*, 292–299.

Huron, D. (2006). *Sweet anticipation*: *Music and the psychology of expectation*. MIT Press.

Jan, S. (2007). *The memetics of music*: *A neo-Darwinian view of musical structure and culture*. Ashgate.

Justus, T. C. & Bharucha, J. J. (2002). Music perception and cognition. In H. Pashler & S. Yantis (Ed.), *Steven's handbook of experimental psychology*: *Sensation and perception* (3rd ed.) (pp. 453–492). John Wiley & Sons Inc.

Kasinathan, V., Mustapha, A., Firdaus, M., Sau Tong, T., & Rahman N. A. A. (2019). Heartbeats: Music recommendation system with fuzzy inference engine. *Indonesian Journal of Electrical Engineering and Computer Science*, *16*, 275–282.

Katz, J., & Pesetsky, D. (2011). *The identity thesis for language and music* [Unpublished manuscript].

Klawonn, F. (2003). Should fuzzy equality and similarity satisfy transitivity? Comments on the paper by M. De Cock and E. Kerre. *Fuzzy Sets and Systems*, *133*(2), 175–180.

Klir, G. J., & Yuan, B. (1995). *Fuzzy sets and fuzzy logic*: *Theory and applications*. Prentice Hall.

Krumhansl, C. (1990). *Cognitive foundations of musical pitch*. Oxford University Press.

Kuo, P. H., Li, T. H. S., Ho, Y. F., & Lin, C. J. (2015). Development of an automatic emotional music accompaniment system by fuzzy logic and adaptive partition evolutionary genetic algorithm. *IEEE Access*, *3*, 815–824.

Kuusi, T. (2001). S*et-class and chord: Examining connection between theoretical resemblance and perceived closeness* [Unpublished PhD thesis]. Sibelius Academy.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.

Leman, M. (1995). *Music and schema theory*. Springer.

Lerdahl, F. (2001). *Tonal pitch space*. Oxford University Press.

Lerdahl, F. (2013). Musical syntax and its relation to linguistic syntax. In M. A. Arbib (Ed.), *Language, music, and the brain*: *A mysterious relationship* (pp. 257–272). MIT Press.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press.

Lerdahl. F., & Krumhansl, C. (2007). Modeling tonal tension. *Music Perception*, *24*(4), 329–366.

Lester, J. (1986). *The rhythms of tonal music*. Southern Illinois University Press.

Lin, C. T., & Lee, C. S. G. (1996). *Neural fuzzy systems*: *A neuro-fuzzy synergism to intelligent systems*. Prentice Hall.

London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press.

Longuet-Higgins, C. H. (1962a). Letter to a musical friend. *Music Review*, *23*, 244–248.

Longuet-Higgins, C. H. (1962b). Second letter to a musical friend. *Music Review*, *23*, 271–280.

Longuet-Higgins, C. H., & Steedman, M. J. (1971). On interpreting Bach. In B. Meltzer & D. Michie (Eds.), *Machine Intelligence 6* (pp. 221–241). Edinburgh University Press.

Lucas, P., Astudillo, E., & Pelàez, E. (2017). Human–machine musical composition in real-time based on emotions through a fuzzy logic approach. *Designing with Computational Intelligence*, *664*, 143–159.

Malcangi, M. (2008). Fuzzy logic-based audio pattern recognition. In T. E. Simos & G. Psihoyios (Eds.), *AIP conference proceedings 1060* (pp. 225–228). American Institute of Physics.

Mamdami, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man–Machine Studies*, *7*(1), 1–13.

Marsden, A. (2010). Schenkerian analysis by computer: A proof of concept. *Journal of New Musical Research*, *39*(3), 269–289.

Marsden, A. (2012). Integrating melodic similarity: A definitive phenomenon or the product of interpretation? *Journal of New Music Research*, *41*(4), 323–335.

Marsden, A., Tojo, S., & Hirata, K. (2018). No longer "somewhat arbitrary": Calculating salience in GTTM-style reduction. In K. Page (Ed.), *DLfM '18: Proceedings of the 5th international conference on digital libraries for musicology* (pp. 26–33). Association for Computing Machinery.

Massaro, D. W., & Cohen, M. M. (1993). The paradigm and the fuzzy logic model of perception are alive and well. *Journal of Experimental Psychology: General*, *122*(1), 115–124.

Mazzola, G. et al. (2002). *The topos of music: Geometric logic of concepts, theory, and performance*. Birkhäuser Verlag.

McKee, E. (2004). Extended anacruses in Mozart's instrumental music. *Theory and Practice*, *29*, 1–37.

Megasari. (2020). Lulunga: Kecapi Suling Instrumental: Rasa Duaan. [Instrumental music]. *YouTube*. https://music.youtube.com/watch?v=XuxzUxeJErw.

Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press.

Milne, A. J. (2013). *A computational model of the cognition of tonality* [Unpublished doctoral thesis]. The Open University.

Mirka, D. (2009). *Metric manipulations in Haydn and Mozart: Chamber music for strings, 1787–1791*. Oxford University Press.

Monti, G., & Sandler, M. (2002). Automatic polyphonic piano note extraction using fuzzy logic in blackboard system. In U. Zölzer (Ed.), *Proceedings of the fifth international conference on digital audio effects* (pp. 39–44). John Wiley & Sons.

Muns, L. (2014). *Music, language, and the deceptive charms of recursive grammars* [Conference session]. *2014 annual conference of the Royal Musical Association Music and Philosophy Study Group*, King's College, London.

Muns, L. (2015). *The inner work of music: Lerdahl and Jackendoff's "Generative theory"* [Unpublished PhD thesis]. University of Amsterdam.

Narmour, E. (1977). *Beyond Schenkerism: The need for alternatives in music analysis*. University of Chicago Press.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures*: *The implication–realization model*. University of Chicago Press.

Norman-Haignere, S. V., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, *88*, 1281–1296.

Novák, V., & Perfilieva, I. (2004). On the semantics of perception-based fuzzy logic deduction. *International Journal of Intelligent Systems*, *19*, 1007–1031.

Nunes–Silva, M., & Haase, V. G. (2013). Amusias and modularity of music processing. *Psychology and Neuroscience*, *6*(1), 45–56.

Oxenham, A. J. (2012). Pitch perception. *The Journal of Neuroscience*, *32*(3), 13335–13338.

Parncutt, R. (1988). Revision of Terhardt's psychoacoustic model of the root(s) of a musical chord. *Music Perception*, *6*(1), 65–94.

Parncutt, R. (1997). A model of the perceptual root(s) of a chord accounting for voicing and prevailing tonality. In M. Leman (Ed.), *Music, gestalt, and computing: Studies in cognitive and systematic musicology* (pp. 181–199). Springer-Verlag.

Parncutt, R. (2011). The tonic as triad: Key profiles as pitch salience profiles of tonic triads. *Music Perception*, *28*(4), 333–365.

Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.

Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, *36*, 767–776.

Paz, I., Nebot, A., Romero, E., & Mugica, F. (2019). Charting perceptual spaces with fuzzy rules. *2019 IEEE international conference on fuzzy systems* (pp. 1–6). IEEE.

Pearce, M. (2016). The group-theoretic description of musical pitch systems [Unpublished manuscript].

Pearce, M., & Wiggins, G. A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, *4*, 625–652.

Peel, J., & Slawson, W. (1984). Review of *A generative theory of tonal music* by Lerdahl, F., & Jackendoff, R. *Journal of Music Theory*, *28*(2), 271–294.

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition*, *100*, 1–32.

Peretz, I. (2009). Music, language and modularity framed in action. *Psychologica Belgica*, *49*(2–3), 157–175.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*(7), 688–691.

Peretz, I., & Zatorre, R. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*, 89–114.

Piston, W. (1941). *Harmony*. W. W. Norton & Co.

Portner, P. H., & Partee, B. (Eds.) (2002). *Formal semantics*: *The essential readings*. Blackwell.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*, 411–440.

Pylyshyn, Z. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*(1–2), 127–158.

Rajendran, V. G., Teki, S., & Schnupp, J. W. H. (2018). Temporal processes in audition: Insights from music. *Neuroscience*, *389*, 4–18.

Rajendran, V. G., Harper, N. S., & Schnupp, J. W. H. (2020). Auditory cortical representation of music favours the perceived beat. *Royal Society Open Science*, *7*, 1–13.

Rameau, J. P. (1971) (1722). *Treatise on harmony*. Dover Publications.

Rawbone, T. (2017). *The butterfly schema as a product of the tendency for congruence and hierarchical selection in the instrumental musical grammar of the classical period* [Unpublished doctoral thesis]. University of Huddersfield.

Rawbone, T. (2021). The conceptual structure of music: Congruence, modularity, and the language of musical thought. In I. D. Khannanov & R. Ruditsa (Eds.), *Proceedings of the world-wide music conference 2021* (pp. 41–54). Springer.

Rawbone, T., & Jan, S. (2020). The butterfly schema in the Classical style: A product of the tendency for congruence. *Music Analysis*, *39*(1), 85–127.

Riemann, H. (1905). *Grundriß der Kompositionslehre* (3rd ed.). Max Hesse.

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

Rosenthal, D. (1992). Emulation of human rhythm perception. *Computer Music Journal*, *16*(1), 64–76.

Ross, T. J. (2010). *Fuzzy logic with engineering applications* (3rd ed.). John Wiley & Sons Ltd.

Rothstein, W. (1989). *Phrase rhythm in tonal music*. Schirmer Books.

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, *112*(29) 8987–8992.

Schenker, H. (1979) (1935). *Free composition* (E. Oster, Tr.). Longman.

Schneider, S. (2011). *The language of thought*: *A new philosophical direction*. MIT Press.

Swain, J. (2002). *Harmonic rhythm*: *Analysis and interpretation*. Oxford University Press.

Suiter, W. (2010). Toward algorithmic composition of expression in music using fuzzy logic. In K. Beilharz, A. Johnston, S. Ferguson, AY-C. Chen (Eds.), *Proceedings of the 2010 conference on new interfaces for musical expression* (pp. 319–322). NIME.

Sun, J., Karray, F., Basir, O., & Kamel, M. (2002). Natural language understanding through fuzzy logic inference and its application to speech recognition. In D. B. Fogel (Ed.), *Proceedings of the 2002 IEEE international conference on fuzzy systems* (pp. 1120–1125). IEEE.

Tavalage, T. M., Sereno, M. I., Melcher, J. R., Ledden, P. J., Rosen, B. R., & Dale, A. M. (2004). Tonotopic organization revealed in human auditory cortex revealed by progressions of frequency sensitivity. *Journal of Neuropsychology*, *91*, 1282–1296.

Temperley, D. (2001). *The cognition of basic musical structures*. MIT Press.

Temperley, D. (2006). *Music and probability*. MIT Press.

Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research*, *1*, 155–182.

Terhardt, E. (1982). Die psychoakustischen Grundlagen der musikalischen Akkordgrundtöne und deren algorithmischen Bestimmung. In C. Dahlhaus & M. Krause (Eds.), *Tiefenstruktur der Musik* (pp. 23–50). Technical University of Berlin.

Todd, N. P. M. (1996). Towards a theory of the principal monaural pathway: Pitch, time and auditory grouping. In S. Ainsworth & S. Greenberg (Eds.), *Proceedings of the workshop on the auditory basis of speech perception* (pp. 216–221). ISCA.

Todd, N. P. M., & Lee, C. S. (2015). The sensory motor theory of rhythm and beat induction 20 years on: A new synthesis and future perspectives. *Frontiers in Human Neuroscience*, *9*, 1–25.

Tojo, S., Marsden, A., & Hirata, K. (2018). On linear algebraic representation of time-span and prolongational reduction trees. In M. Aramaki, M. Davies, R. Kronland-Martinet, & S. Ystad (Eds.), *Music technology with swing, CMMR 2017*: *Lecture notes in computer science* (pp. 199–212). Springer.

Tokumaru, M., Yamashita, K., Muranaka, N., & Imanishi, S. (1998). Membership functions in automatic harmonization system. In B. Werner (Ed.), *Proceedings* of the *1998 28th IEEE international symposium on multiple-valued logic* (pp. 350–355). IEEE.

Tymoczko, D. (2012). *The geometry of music: Harmony and counterpoint in the extended common practice*. Oxford University Press.

Tymoczko, D. (2022). Hierarchical set theory. *Journal of Mathematics and Music*, *17*(2), 282–290.

Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences* [Unpublished PhD dissertation]. Eindhoven University of Technology.

Weyde, T., & Dalinghaus, K. (2001). Recognition of musical rhythm patterns based on a neuro-fuzzy system. In C. H. Dagli et al. (Eds.), *Smart engineering system design: Neural networks, fuzzy logic, evolutionary programming, data mining, and complex systems* (pp. 679–684). ASME Press.

Yahia, B. N., Bellamine, N., & Ghezala, H. B. (2012). Integrating fuzzy case-based reasoning and particle swarm optimization to support decision making. *International Journal of Computer Science Issues*, *9*(3), 117–124.

Yang, Y. H., Liu, C. C., & Chen, H. H. (2006). Music emotion classification: A fuzzy approach. In K. Nahrstedt et al. (Eds.), *Proceedings of the 14th annual ACM international conference on multimedia* (pp. 81–84). ACM.

Yeh, R. T. (1973). Toward an algebraic theory of fuzzy relational systems. *Proceedings of the international congress on cybernetics* (pp. 205–223). IEEE.

Yilmaz, A. E., & Telatar, Z. (2010a). Fuzzy logic based four-voice choral harmonisation in traditional style. *Journal of Intelligent and Fuzzy Systems*, *21*, 289–301.

Yilmaz, A. E., & Telatar, Z. (2010b). Note-against-note two-voice counterpoint by means of fuzzy logic. *Knowledge Based Systems*, *23*, 256–266.

Yost, W. A. (2009). Pitch perception. *Attention, Perception, & Psychophysics*, *71*(8), 1701–1715.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.

Zadeh, L. A. (1971). Similarity relations and fuzzy ordering. *Information Sciences*, *3*, 177–200.

Zatorre, R. J., Belin, P., & Penhume, V. B. (2002). Structure and function of auditory cortex. *Trends in Cognitive Sciences*, *6*(1), 37–46.

Zhang, H., & Liu, D. (2006). *Fuzzy modeling and fuzzy control*. Birkhäuser.