

A Cocreative Computational Approach to Musical Analogy

Abstract

Keywords

Description and Goals

Analogy at the Core of Musical Thought

Structure-Mapping and Analogy in Artificial Intelligence

Audio Synthesis Through Analogy-Driven Mappings

References

A Cocreative Computational Approach to Musical Analogy

Nuno Trocado ^{1,2},

¹ CEIS20 – Center for Interdisciplinary Studies, University of Coimbra, Portugal

² Faculty of Arts and Humanities, University of Coimbra, Portugal

nuno@nunotrocado.com

Abstract

Analogy is a critical cognitive process, at the core of the multiple ways in which we think *in and through* music. With structure-mapping theory as a point of departure, I describe how a computational implementation of its theoretical tenets may frame an approach to musical analogy which, in a cross-domain or music-to-music generative setup, would amount to a novel variation of concatenative synthesis, but driven preferably by higher-order relational structures instead of by the mere similarity of feature vectors.

Keywords

Analogy, Music, Synthesis, Computational creativity

Description and Goals

In this paper, I describe a hypothetical technological framework for sound synthesis, which works as the other half of a human-machine cocreative system, grounded at least in part on a simulation of the cognitive capacity for analogy-making. My goal is not to design a system that generates music integrally or autonomously—i.e., where the machine appears to act by its own accord, producing finished or almost finished works or segments of works—but to better understand analogy and how it can be computationally exploited to foster and enhance human creativity, in particular in the context of my own artistic practice and the aesthetic values that it entails.

Analogy at the Core of Musical Thought

Through analogy we compare things that are different, but share relevant commonalities, allowing us to project cognitively resonant structures between them and gain new functional insights. As the “fuel and fire of thinking” (Hofstadter and Sander 2013), analogy is central to a wide range of human abilities, ubiquitous in everyday thought, and determinant for our worldly experience. It is thus unsurprising that analogy shows up prominently in music. Participating in the musical phenomenon involves the cognition of sound-pattern formation and the mapping of gestural-temporal processes, which shape how we make, listen, think in or about music, feel and move through it, individually and collectively. Knowledge from a variety of domains is carried over, or projected, into sound, thus constituting our musical experience. Such projection is the characteristic mark of analogy.

Notwithstanding the centrality of analogical processes, musical studies have only recently started to examine their conceptual implications. Some accounts focus on correspondences through recurrence “within music,” such as thematic/formal roles (Bourne 2015), where the schematic repetition or transformation of a pattern in successive musical passages gives shape to “chains of analogy” (Kielian-Gilbert 1990). As analogical comparisons drive the processes of conceptualization and abstraction, analogy lies at the core, for example, of the concept of *motives*, through which we understand each new instance of a musical pattern by comparing it with other ones that we previously encountered, noting their shared structure despite the superficial dissimilarities. Analogy is also implicated in the cognition of metric groupings, or the exposition and recapitulation of a sonata form, or for that matter in the chorus/verse recurrence of a pop song. Other accounts focus on structural commonalities between different parameters, e.g., pitch and time (Bar-Yosef 2007; Eitan and Granot 2007). Furthermore, the suggestion that music works fundamentally as a “sonic analog for dynamic processes” (Zbikowski 2017) connects music with emotion, gesture, dance, or words.

These connections mount to the view that our conceptual system is prominently *metaphoric*, i.e. constrained by the features of the human body and worldly experience (Lakoff and Johnson 1980), and structured by sensorimotor schematic patterns (Johnson 2007). Such a perspective, together with the analogy-like theory of *conceptual blending* (Fauconnier and Turner 2002), prompted emergent frameworks on how music is conceptualized (Brower 2000; Hatten 1995; Larson 2012; Saslaw 1996; Spitzer 2004; Zbikowski 2002).

Since, in these cases, music appears to be “standing for” a distinct reality, Zbikowski (2017) highlights that music makes use of a unique form of reference—“analogical reference”—which can be understood in terms of Peircean semiotics, and particularly in terms of the concept of the *icon*. Zbikowski observes that, as Peirce divided the icon into *image*, *diagram*, and *metaphor*, sonic analogs can be traced in a continuum between those categories, where sounds that more clearly mimic an actual audible event are positioned closer to image, and the sonic analogs for nonsonic dynamic processes closer to metaphor. *Symbolic* reference, by contrast, while predominant in language, is residual in music, being relegated to instances where a musical utterance is *conventionally* correlated with a specific referent, as it notably happens with the culturally shared associations that constitute the object of topic theory.

Structure-Mapping and Analogy in Artificial Intelligence

The computational modeling of analogy enjoys a rich history. Understandably, if the capacity for analogy is such a critical mark of intelligence, it follows that it must be somehow introduced in artificial intelligence systems. On the other hand, research on analogy as a cognitive process arose and was developed contemporaneously with the general perspective positing that human reasoning can be understood through its implementation in computer programs. The technical approaches for artificial analogy-making have followed the broader trends in the field of artificial intelligence, from older (but still promising) symbolic methods, which are based on the manipulation of symbols representing the knowledge for the base and target domains, to more recent deep learning techniques, as well as hybrid architectures. Reviewing several of these approaches, Mitchell (2021) concludes that despite the extensive efforts, which remain as active as ever, “no current AI system is anywhere close to a capability of forming humanlike abstractions or analogies,” while at the same time, such advances will be key for continued progress going forth from current state-of-the-art artificial intelligence models.

According to structure-mapping theory (Gentner 1983, 1989; Gentner and Smith 2013; Gentner et al. 2001), developed for the last four decades and now in some ways the classic, empirically validated framework for analogy, we are biased toward mapping relational structures, and preferably systems of mutually connected higher-order relations, and not so much object properties or attributes—this preference is called the *systematicity principle*. This is why we find the analogy between a house and a nest (same functional relationships) more compelling than the one between a planet and a ball (same shape)—see Figure 1.

A house is like a nest. A ball is like a planet.

Common interconnected relations: contains a creature offers protection from threats built for that purpose protection extends to fragile offspring center of family life

Common object properties: roundness

Figure 01. The house/nest analogy is more compelling than the ball/planet one.

Through this distinction between relations and properties or attributes, it’s possible to contrast (as in a continuum rather than rigorously separated categories) analogy with other types of domain comparison (see Table 1). In *literal similarity* the mapping includes a large number of both object attributes and relationships. *Mere appearance* hinges on common attributes, but not relations. In *abstraction*, as in analogy, there are few attributes mapped to the target, but the base domain is already an abstract relational structure, which has few (or none) object attributes to begin with. Finally, a comparison presenting neither attribute nor relational overlap is an *anomaly*.

Table 01. Analogy among other types of domain comparison. Reproduced from Gentner (1989).

	Attributes	Relations	Example
Literal similarity	Many	Many	Milk is like water
Analogy	Few	Many	Heat is like water
Abstraction	Few	Many	Heat flow is a through-variable
Anomaly	Few	Few	Coffee is like the solar system
Mere appearance	Many	Few	The glass tabletop gleamed like water

The Structure-Mapping Engine, SME (see Forbus et al. 2016 for current iteration), is a computational implementation of the structure-mapping theory. Like comparable systems, it follows significant assumptions: that analogy is domain-general, its mechanism is purely syntactic and not constrained by the specific perceptual modes involved in the process; and that therefore it is independent of the way through which knowledge is structured in the base and target domains. Analogy is thus described as a neutral mechanism, operating in the same fundamental way between domains like water and heat, as between two sets of different geometric drawings, or, say, between sound and the kinesthetic patterns evidenced in dancing. This means that, as a different and previous step to the analogy proper, there is the need for construing explicit domain-knowledge representations, in particular representations that go beyond flat feature vectors and capture nth-order relational structures—i.e., that designate the relations (and relations between relations) drawing up the structural constituents of the domain. These representations, however, don’t have to be hand-coded and can be automatically generated or derived from perceptual input.

Computational models such as the SME are in a sense disembodied, but it can be argued that they remain compatible with a connection tracing back domain knowledge to its roots in modality-specific, sensorimotor representations. Furthermore, the encoding and matching modules, while independent, can be interleaved, feeding back into each other—this process mirrors the empirical observation that high-level cognitive processes penetrate into and affect the operation of perceptual systems. Additionally, human intelligence and creativity may indeed be impossible to simulate on a full-scale model, or even in a less-ambitious imperfect simulacrum, as they are dynamically contingent on the features and history of the body, intertwined with environmental factors, and dependent on the specific, more or less unpredictable goals pursued by the agent. But, even if machines don’t possess these things, simulated outcomes remain pragmatically useful, either as a heuristic—furthering partial accomplishments and a more profound understanding of human cognition—or, in the sense that most concerns this endeavor, as an aesthetically valuable instrument for artistic practice, that retains a solid connection to the psychology of musical experience.

Audio Synthesis Through Analogy-Driven Mappings

The idea of applying cognitively-resonant domain-general computational models of analogy to music, or integrating implementations such as the SME to sound generation tasks, remains largely unexplored. Some tentative music-related approaches (Eppé et al. 2018; Zacharakis et al. 2021) have instead followed the *conceptual blending* framework (Fauconnier and Turner 2002), which describes a very similar high-level cognitive process where elements and relations from two or more domains are compared, but conceptualizing their combination as a fusion (*blend*) into a new integrated mental space. The integration network model is meticulously specified and, as it is apt for formalization, has been an attractive framework for computational approaches. On the other hand, conceptual blending is targeted to the creation of hybrids, and thus it’s less flexible than more general models of analogy.

Otherwise, I find the tenets of structure-mapping particularly apt for the domain of music. Music is highly relational, at the very least because of its intrinsic temporality. What is the value of a single sound event, if it’s not taken in relationship with past and future ones? Besides, higher-order relations are manifest in the pervasiveness of conceptualizations that organize sound in hierarchies, processual configurations, or cause-effect chains.

In this context, I envision a representation of the sound domain that proceeds from segmenting an audio stream into multiple very short sonic tokens, quantified according to music information retrieval metrics, which are correlated temporally with image schematic patterns such as containment, source-path-goal, interruption, self-similarity, or pendulum. Such information is then stored in a corpus database. From here, the SME probes and acts upon hypothetical cross- or intra-domain mappings. Cross-domain mappings are made possible by having the non-sonic domain categorized through the same common image schemas. Music-to-music mappings would have another audio stream as the base for the analogy.

The creation of new sounds is accomplished through a kind of concatenative synthesis (Schwarz 2004)—a method of generating audio by selecting and assembling small sonic units from a large database of sound sources. Typically, the selection and assemblage are performed by attempting to match quantitative physical, perceptual, or statistical features (e.g., pitch, spectral centroid, average amplitude, tempo) of the sources. Such feature-matching depends on a specification of criteria for the similarity between sonic units. In the various kinds of domain comparison that were contrasted above, this kind of similarity would approximate “mere appearance,” since it deals predominantly with collections of object properties. In the analogy-driven setup that I propose, however, mappings would be established not according to the similarity of surface features, but according to the degree of isomorphism in relational structures.

Thus, mapped sonic units would not necessarily sound similar; instead, the resulting audio stream would exhibit deeper structural commonalities perceived as convincing, compelling, and surprising, despite the superficial differences—just like the analogies that we rely upon in our day-to-day life.

In conclusion, I believe that this strategy leads to a machine-generating but human-steerable framework for producing novel timbres and sonic textures. One that, by being grounded in the cognitive capacity for analogy, exhibits a degree of creativity still lacking in artificial intelligence systems, and whose expected glitches, non-linearities, and incoherences could be artistically useful in music-making.

References

- Bar-Yosef, Amatzia. 2007. “A Cross-Cultural Structural Analogy Between Pitch and Time Organizations.” *Bourne Perception* 24 (3): 265–80. <https://doi.org/10.1525/mp.2007.24.3.265>.
- Bourne, Janet. 2015. “A Theory of Analogy for Musical Sense-Making and Categorization: Understanding Musical Jabberwocky.” PhD thesis, Northwestern University.
- Brower, Candace. 2007. “A Cognitive Theory of Musical Meaning.” *Journal of Music Theory* 44 (2): 323. <https://doi.org/10.2307/3090681>.
- Eitan, Zohar, and Roni Y. Granot. 2007. “Intensity Changes and Perceived Similarity: Inter-Parametric Analogies.” *Musicae Scientiae* 11 (1): 39–75. <https://doi.org/10.1177/1029864907011001031>.
- Eppé, Manfred, Ewen Maclean, Roberto Confalonieri, Oliver Wolff, and Consuelo Boronat. 2001. “Metaphor Is Like Analogy.” In *The Analogical Mind: Perspectives from Cognitive Science*, edited by D. Gentner, K. J. Holyoak, and B. N. Kokinov, 199–253. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/1251.003.0010>.
- Gentner, Dedre, and Linsey A. Smith. 2013. “Analogical Learning and Reasoning.” In *The Oxford Handbook of Cognitive Psychology*, edited by Daniel Reisberg. New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195376746.013.0042>.
- Hatten, Robert S. 1995. “Metaphor in Music.” In *Musical Signification*, edited by Eero Tarasti, 373–92. De Gruyter Mouton. <https://doi.org/10.1515/9783110885187.373>.
- Hofstadter, Douglas, and Emmanuel Sander. 2013. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*. New York: Basic Books.
- Johnson, Mark. 2007. *The Meaning of the Body*. Chicago: University of Chicago Press.
- Kielian-Gilbert, Marianne. 1990. “Interpreting Musical Analogy: From Rhetorical Device to Perceptual Process.” *Music Perception* 8 (1): 63–94. <https://doi.org/10.2307/40285486>.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live by*. Chicago: University of Chicago Press.
- Larson, Steve. 2012. *Musical Forces: Motion, Metaphor, and Meaning in Music*. Bloomington, MN: Indiana University Press.
- Mitchell, Melanie. 2021. “Abstraction and Analogy-Making in Artificial Intelligence.” *Annals of the New York Academy of Sciences* 1505 (1): 79–101. <https://doi.org/10.1111/nyas.14619>.
- Saslaw, Janina. 1996. “Forces, and Paths: The Role of Musical Sense-Making and Derived Image Schemas in the Conceptualization of Music.” *Journal of Music Theory* 40 (2): 217. <https://doi.org/10.2307/843889>.
- Schwarz, Diemo. 2004. “Data-Driven Concatenative Sound Synthesis.” PhD thesis, Université Paris 6 – Pierre et Marie Curie. <http://recherche.ircam.fr/equipements/analyse-synthese/schwarz/thesis/>.
- Spitzer, Michael. 2004. *Metaphor and Musical Thought*. Chicago: University of Chicago Press.
- Zacharakis, Asterios, Maximos Kaliakatsos-Papakostas, Stamatia Kalaitzidou, and Emilios Cambouropoulos. 2021. “Evaluating Human-Computer Co-Creative Processes in Music: A Case Study on the Chameleon Melodic Harmonizer.” *Frontiers in Psychology* 12. <https://doi.org/10.3389/fpsyg.2021.603752>.
- Zbikowski, Lawrence M. 2002. *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. Oxford: Oxford University Press.
- . 2017. *Foundations of Musical Grammar*. New York: Oxford University Press. <https://doi.org/10.1093/oso/9780190653637.001.0001>.