Problems with Cash Rewards

Issues and Music and Sciences

Dr. David John Baker
HU Berlin, Winter 2020

Preface

Outline

I. Some problems that music science academics work on also have industry application
II. Today (and next lecture) we take long tour of types of problems where answer comes with cash reward
III. Though engineering based, solutions here depend on domain knowledge of music!
IV. We of course stop to note any issues with each problem

Music as Complex Behavior

Music as Complex Behavior

- Much of course content has been about investigating questions for the sake of understanding
    - Answering the big questions!
    - What is music?
    - How do animals perceive what humans call music?
    - What does it mean for a question about music to be scientific?
    - All assert that music can be understood as complex human activity
- What we do in academia can and does have relationship to the way the rest of the world functions (capitalism!)

Music as Complex Behavior

- In addition to generating ideas and new ways of thinking, researchers also apply ideas from academia to business settings
- If you know how world works → Exploit for your benefit
    - Can be done for good
    - Can be done for evil
    - Is being done and is important to question and critique (what we have luxury of doing right now !!)
- Show several examples of what happens when academic thought is translated into goods and services

Whistle Stop Tour

I. Audio Branding
II. Music and Advertising
III. Behavioral Economics of Music
IV. Audio Fingerprinting
V. Music Generation
VI. Recommendation Systems

Audio Branding

Audio Branding

If you had to describe Apple (the company) as a person, what would they be like? What traits would you associate with the person? What traits do they not embody?

Audio Branding

If you had to describe Apple (the company) as a person, what would they be like? What traits would you associate with the person? What traits do they not embody?

Where do these feeling come from?

How would Meyer describe how this works?

Audio Branding

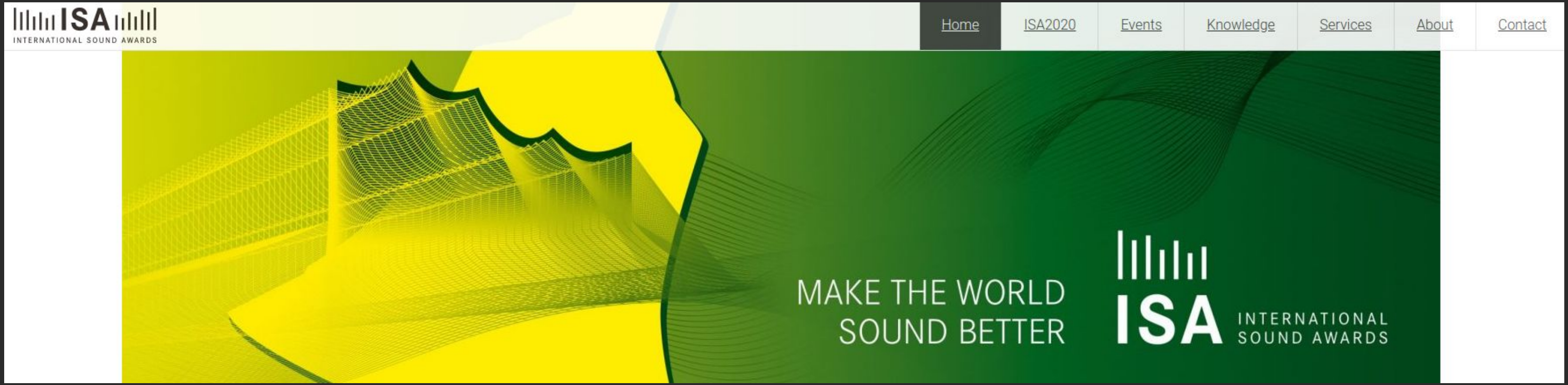
1. If you heard this music in an advertisement, what type of product do you think would be being sold?
2. What group of people might be the target demographic?
3. How can we know if our intuitions are correct?
4. How is this similar to the visual brand personality from before?

Audio Branding

What Is Audio Branding?

Music and advertising – a combination that causes some people to associate fantasies of omnipotence, and causes others to be afraid of unwanted manipulation. What is behind all this? Are there hidden persuaders and subliminal advertising? Or is it "only" about emotionalizing brands?

Audio Branding describes the process of brand development and brand management by use of audible elements within the framework of brand communication. It is part of multi-sensory brand communication and holistic brand design. Audio Branding aims at building solidly a brand sound that represents the identity and values of a brand in a distinctive manner. The audio logo, branded functional sounds, brand music or the brand voice are characteristic elements of Audio Branding.



Audio Branding

- Represents Identity of the Brand
- Often associated with Audio Logo
- Can also extend to functional sounds
    - Turn on computer, alert medical staff in hospital
- Music associated with Brand
- Voices associated with Brand
    - (who is doing the voice over on the commercial, what does that convey?!)

How do we formalize these intuitions about sound and identity?

Audio Branding

- Need systematic way to talk about the complexities associated with music, sound, culture, audio
- Creatives team with data scientists to come up with ideas that can be tested
- Often employs focus groups or pre-rating
  - Focus Groups: Bring group of people together that are representative of target demographic, ask for opinion on new products
  - Pre-testing: Create several possible versions of audio, marketing, media campaign and see in small group which does best
- Need scientific methods in order to avoid confounds

Audio Brading

Why would it be important to have representative sampling in focus groups?

Why could you assume that a track that did well on a pre-test would also succeed in a multi-million dollar campaign?

Music and Advertising

Music and Advertising

- Music frequently used in advertising campaigns
- Used also to set mood / vibe in other settings
- Active Area of Research of Music + Background
  - Garlin and Owen 2006
  - Familiarity and liking have positive effect on patronage
  - Presence of music has positive effect on consumers
  - Slower tempos associated with customers staying longer
  - Higher volume and tempo leads to longer time duration perception
  - Tempo has greatest effect on arousal
- Theoretical constructs used to help understand why music would elicit these effects

What consumer situations would you want to make time feel faster? slower?

Congruence Matching

- North, Hargreaves, and McKendrick (1997, 1999a)
- Priming paradigm ** (Highly unstable area of research)
- Wine store
  - Have both German and French Wine
  - Put tape player (the 90s!) near wine
  - French music → Increased sale of french wine
  - German music → Increased sale of German wine
  - Though to be some sort of congruency effect
- External Hiring at big companies

Melodic Memorability

- Music is effective way to retain information
  - Singing ABCs
  - Audio logo uniqueness / Jingles
- Overlap in recall/recognition of melodic memory literature
- Use symbolic features of songs to dissect what makes song easy/difficult to remember
- Overlap with earworm literature
  - Song stuck in Head
- Jakubowski et. al 2017 → Dissecting Earworm

Melodic Memorability

Psychology of Aesthetics, Creativity, and the Arts
2017, Vol. 11, No. 2, 122–135

1931-3896/17/$12.00 http://dx.doi.org/10.1037/aca0000090

Dissecting an Earworm: Melodic Features and Song Popularity Predict Involuntary Musical Imagery

Kelly Jakubowski
Goldsmiths, University of London

Sebastian Finkel
University of Tübingen

Lauren Stewart
Goldsmiths, University of London and Aarhus University and The Royal Academy of Music Aarhus/Aalborg, Denmark

Daniel Müllensiefen
Goldsmiths, University of London

Involuntary musical imagery (INMI or "earworms")— the spontaneous recall and repeating of a tune in one's mind—can be attributed to a wide range of triggers, including memory associations and recent musical

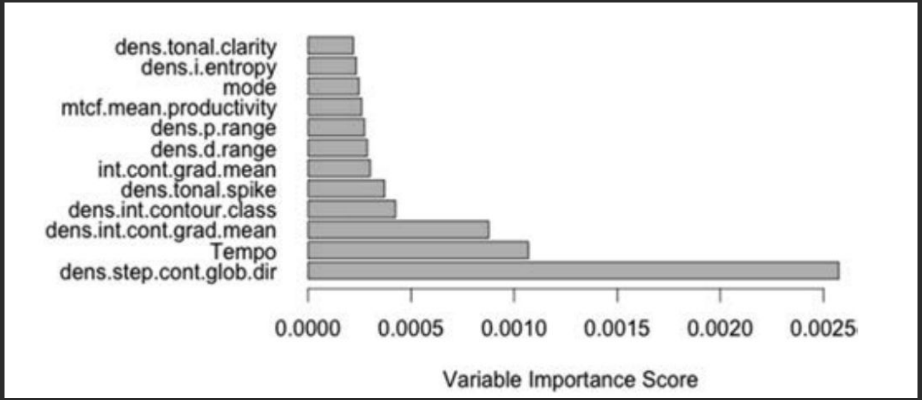
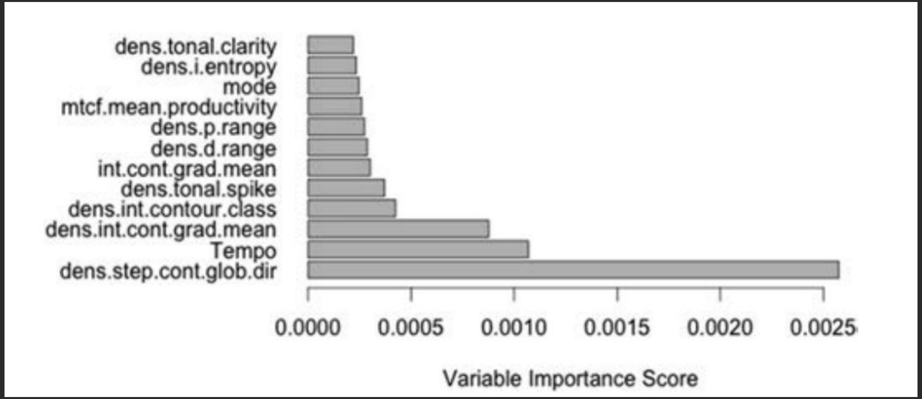


Table 1
Songs Most Frequently Named as Involuntary Musical Imagery (INMI)

Song title, artist	Number of times named as INMI	Highest chart position	Weeks in charts	Days since chart exit
(1) "Bad Romance," Lady Gaga	33	1	47	1,322
(2) "Can't Get You Out of My Head," Kylie Minogue	24	1	25	4,164
(3) "Don't Stop Believing," Journey	21	6	59	1,399
(4) "Somebody That I Used to Know," Gotye	19	1	46	398
(5) "Moves Like Jagger," Maroon 5	17	2	52	545
(6) "California Gurls," Katy Perry	15	1	26	1,083
(7) "Bohemian Rhapsody," Queen	14	1	17	13,621
(8) "Alejandro," Lady Gaga	12	7	10	1,175
(9) "Poker Face," Lady Gaga	11	1	66	1,490

Melodic Memorability

Figure 4. Examples from the present dataset with the (A) lowest and (B) highest values of the variable *dens.step.cont.glob.dir*. INMI = involuntary musical imagery.



Behavioral Economics of Music

Homo Economicus

- Assumption that humans are
  - Rational actors (do what makes sense logically)
  - Self interested (just think about what benefits self)
  - Look for “optimal solutions”
- Compute and maximize utility function
  - Always pick what makes most sense given options available
- Example: Two cars
  - Both have clear function (Point A → Point B)
  - Clear difference in price
  - Porsche = 80K GBP
  - 1983 Oldsmobile Cutlass Ciera = 8K GBP
- Why pick Porsche over the Cutlass?

In Rainbows

- 2007 Radiohead Releases In Rainbows
- Initially released as digital download
- Fans “pay what you wish”
- Box set later this year

What explains why everyone didn’t just download it for free?

THE DECADE IN MUSIC: '00S

The 'In Rainbows' Experiment: Did It Work?

November 16, 2009 · 10:00 PM ET

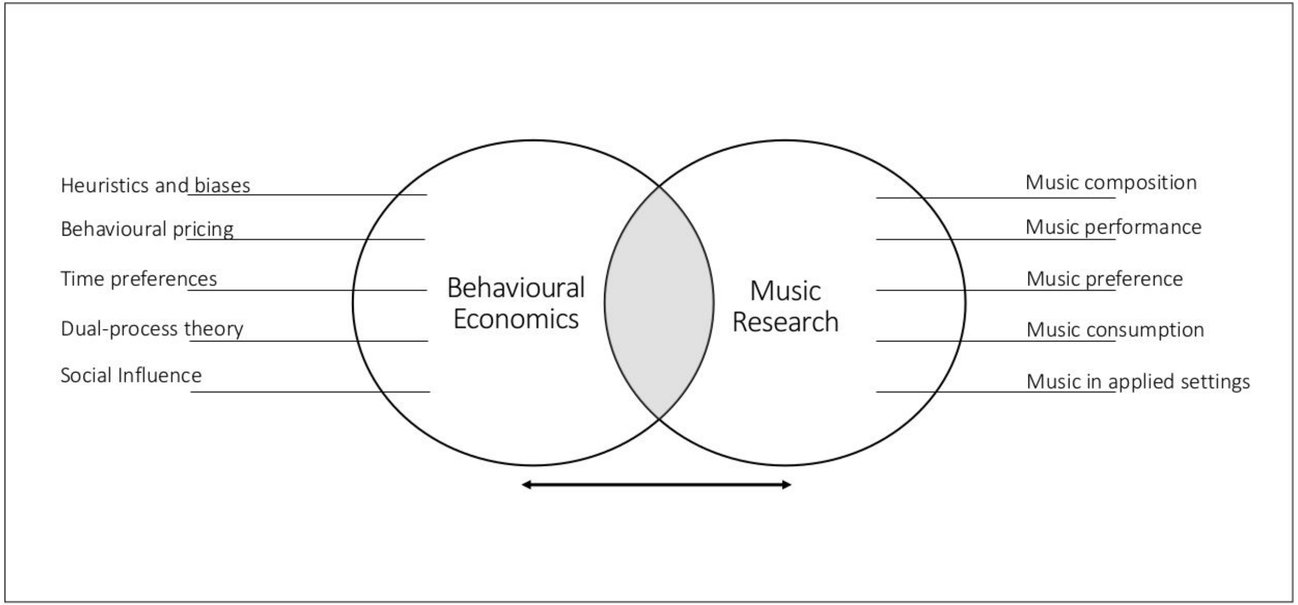
ERIC GARLAND

In 2007, Radiohead no longer had a contract with its record company, EMI, and was about to release its seventh full-length album, *In Rainbows*. For its first new release in four years, one of the world's biggest bands would have no major-label representation.

In Rainbows was first released on Oct. 10, 2007, as a digital download from the band's Web site. Fans were encouraged to "pay what you wish" — even nothing — and a "digital tip jar" was set up to collect voluntary payments. On Dec. 4 of that same year, an $80 deluxe box set

Music as the Product (Anglada-Tort, 2021)

Fig. 1.1 The Behavioural Economics of Music.



Music and Consumer Decision Making

- Behavioral Economics of Music
  - Largely inspired by work of Kahneman and Tversky → Thinking Fast and Slow
  - Important for inspiring new ideas, not as stable findings since initial publication
- Examples
  - Repeated Song Illusion
  - Linguistic Fluency

Table 2.2 Heuristics and biases identified in the systematic review (S2).

	Definition	Music Example
Processing fluency	Human tendency to evaluate easy-to-process information more positively than similar but more difficult-to-process information (Reber et al., 2004).	Songs with more repetitive lyrics, which are easier to process in terms of information, are perceived as being liked more (Nunes et al., 2015).
Availability heuristic	When judging the frequency and probability of events, people rely on the ease with which examples come to their minds (Tversky & Kahneman, 1974).	Listeners falsely remember sounds that come easily to their minds (Vuvan et al., 2014).
Representativeness heuristic	People estimate the likelihood of an event by comparing it to an existing event of similar characteristics that already exists in their minds (Tversky & Kahneman, 1974).	Stereotypes between music genres and fans can be misjudged (Lonsdale & North, 2011).
Affect heuristic	Human tendency to rely heavily upon our emotional state when making judgements and decisions (Slovic et al., 2002).	Individuals evaluate music based on an associated emotional feeling (Anglada-Tort et al., 2018)
Framing effect	People make decisions based on how the options are presented or "framed" (e.g., as a loss or as a gain) (Kahneman & Tversky, 1979).	Contextual information presented with music can systematically affect a person's judgement of the music (North & Hargreaves, 2004)
Peak-end rule	People judge an experience largely based on how they felt at its peak (i.e., the most intense point) and at its end (Kahneman & Fredrickson, 1993).	Listeners evaluate a music experience based on the most intense moment and at the end (Rozin, et al., 2004).

Music and Consumer Decision Making

- Behavioral Economics of Music
  - Largely inspired by work of Kahneman and Tversky → Thinking Fast and Slow
  - Important for inspiring
    stable findings since i
- Examples
  - Repeated Song Illusion
  - Linguistic Fluency

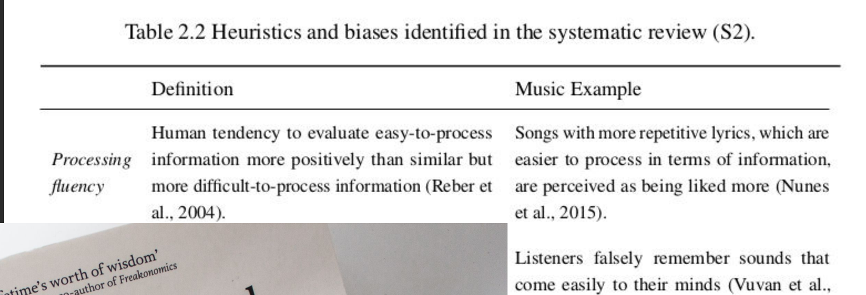
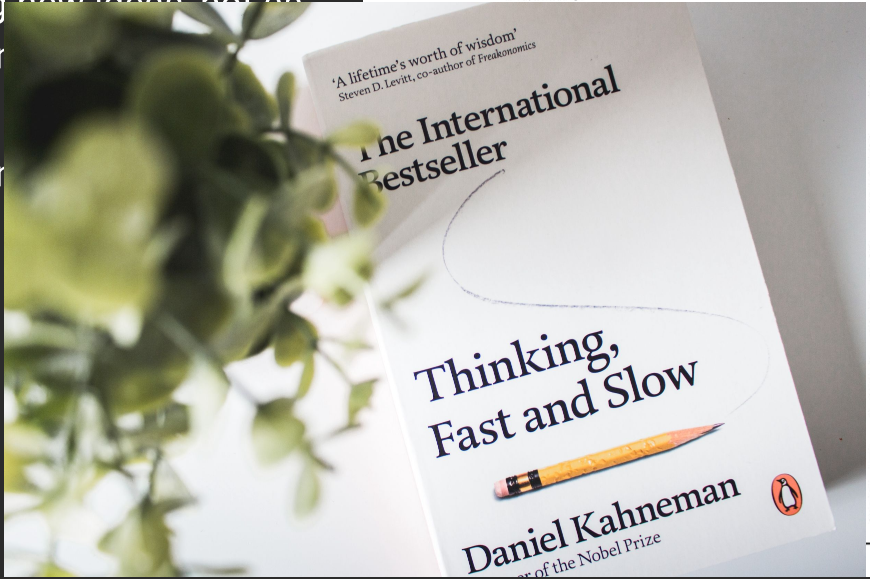


Table 2.2 Heuristics and biases identified in the systematic review (S2).

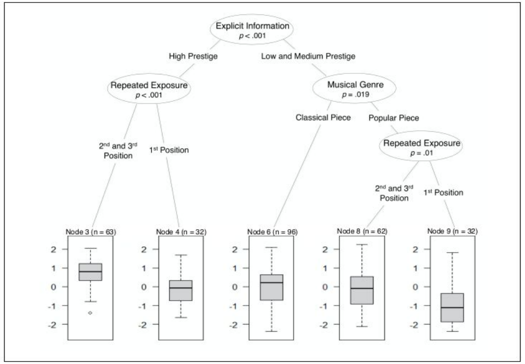
	Definition	Music Example
Processing fluency	Human tendency to evaluate easy-to-process information more positively than similar but more difficult-to-process information (Reber et al., 2004).	Songs with more repetitive lyrics, which are easier to process in terms of information, are perceived as being liked more (Nunes et al., 2015).
		Listeners falsely remember sounds that come easily to their minds (Vuvan et al., 2014).
		Stereotypes between music genres and fans can be misjudged (Lonsdale & North, 2011).
		Individuals evaluate music based on an associated emotional feeling (Anglada-Tort et al., 2018)
		Contextual information presented with music can systematically affect a person's judgement of the music (North & Hargreaves, 2004)
		Listeners evaluate a music experience based on the most intense moment and at the end (Rozin, et al., 2004).



Repeated Song Illusion

- Participants heard three songs
- Each song described by “different” songs and meta-information
  - Low
  - Medium
  - High Prestige
- In reality, songs were the same
- 75% of people believed they heard different songs
- Prestige INCREASES perception of several musical factors

Fig. 2.2 The influence of non-musical factors on music performance evaluation.



The tree model is useful in identifying the most predictive variables influencing music performance evaluation, as well as specific conditions that lead to particularly high (left node, 3) and low (right node, 9) ratings. The tree model can be interpreted by starting at the top and following each branch down, to arrive at a terminal node. A path to a terminal node describes the interaction of experimental conditions that lead to a particular subset of ratings.

Anglada-Tort, M., & Müllensiefen, D. (2017). The repeated recording illusion: the effects of extrinsic and individual difference factors on musical judgments. *Music Perception: An Interdisciplinary Journal*, 35(1), 94-117. https://doi.org/10.1525/mp.2017.35.1.94

Linguistic Fluency

- Participants presented both easy and difficult to pronounce song titles
- (Turkish words, to native English speakers)
- More positive associations with titles that were easier to pronounce
- What else could have been driving this effect?

Table 1

Fluent and Disfluent Turkish Names

Fluent	**Disfluent**
Dermod by Artan	*Siirt* by Lasiea
Kado by Pera	*Taahhut* by Aklale
Boya by Tatra	*Emniyet* by Luici
Alet by Ferka	*Dizayn* by Sampiy

BEM Summary

- Models of “rational” thought and consumer behaviour not completely explained by homo economicus
- Music, as example of complex behavior, is great way to study human behavior
- Music can be used as way to understand complex behavior
- Humans susceptible to extra-musical information

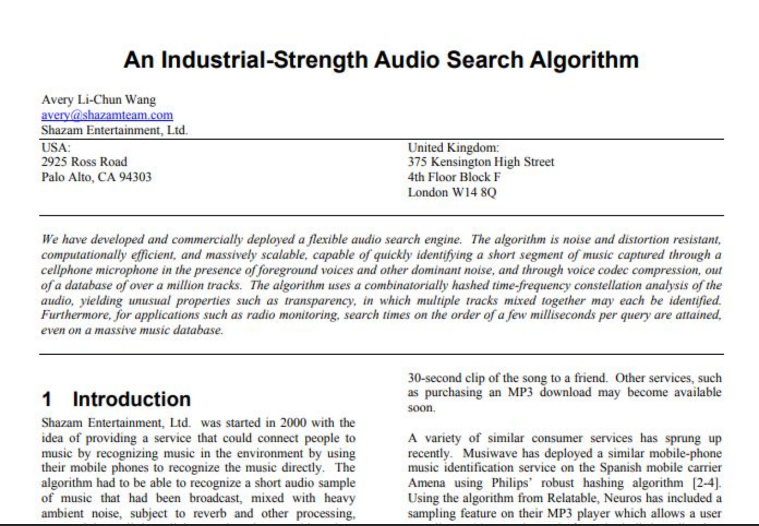
Audio Fingerprinting

Audio Fingerprinting

- Learned before that music can be conceptualized as a product
    - People want to buy it
    - Also cultural currency to have songs in your playlist, discussions
- If that is the case, how do you find out what a song is if the DJ doesn’t say the name of it? You’re in loud coffee shop?
- Shazam !!
- How does it work!?

Shazam Algorithm

- Find way to search for musical metadata (name of song, artist) with audio
- Prereq:
  - Need database of songs it could be…
  - Need way to match incoming audio sample with database
- Extract salient pitch structures (represented in spectrogram)
- https://musiclab.chromeexperiments.com/Spectrogram/
- Match with larger database



An Industrial-Strength Audio Search Algorithm

Avery Li-Chun Wang
avery@shazamteam.com
Shazam Entertainment, Ltd.

USA:
2925 Ross Road
Palo Alto, CA 94303

United Kingdom:
375 Kensington High Street
4th Floor Block F
London W14 8Q

We have developed and commercially deployed a flexible audio search engine. The algorithm is noise and distortion resistant, computationally efficient, and massively scalable, capable of quickly identifying a short segment of music captured through a cellphone microphone in the presence of foreground voices and other dominant noise, and through voice codec compression, out of a database of over a million tracks. The algorithm uses a combinatorially hashed time-frequency constellation analysis of the audio, yielding unusual properties such as transparency, in which multiple tracks mixed together may each be identified. Furthermore, for applications such as radio monitoring, search times on the order of a few milliseconds per query are attained, even on a massive music database.

1 Introduction

Shazam Entertainment, Ltd. was started in 2000 with the idea of providing a service that could connect people to music by recognizing music in the environment by using their mobile phones to recognize the music directly. The algorithm had to be able to recognize a short audio sample of music that had been broadcast, mixed with heavy ambient noise, subject to reverb and other processing,

30-second clip of the song to a friend. Other services, such as purchasing an MP3 download may become available soon.

A variety of similar consumer services has sprung up recently. Musiwave has deployed a similar mobile-phone music identification service on the Spanish mobile carrier Amena using Philips' robust hashing algorithm [2-4]. Using the algorithm from Relatable, Neuros has included a sampling feature on their MP3 player which allows a user

Shazam Algorithm

- 1A: Capture spectrogram
  - Audio in frequency domain
- AB: Extract salient pitch markers
- From anchor point, find other matches
- Compute against big table of other matches
- Not operating on scores!
- Operating on representation of music/audio that is available to humans but we don't conceptualize as!
- https://youtu.be/WhXgpkQ8E-Q

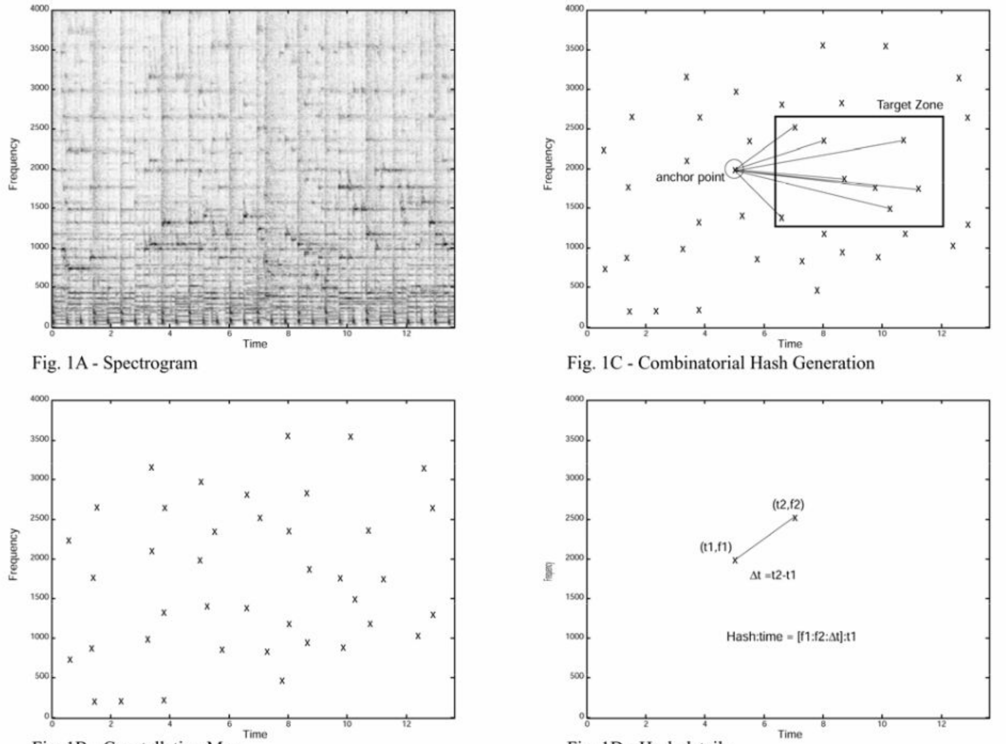


Fig. 1A - Spectrogram

Fig. 1C - Combinatorial Hash Generation

Fig. 1B - Constellation Map

Fig. 1D - Hash details

Music Generation

Artificial Intelligence and Music

- First ever AI Google Doodle
- Enter melody on staff → Harmonized
- Breaking it down
  - Create representational system
  - Teach (train) AI on existing Data
  - Learn Patterns
  - Reproduce Summary of Patterns
- Read more
  - https://magenta.tensorflow.org/coconet

What are the important elements of music that a computer would need to be able to understand in order to represent the sonic elements of music?

How could you do this?

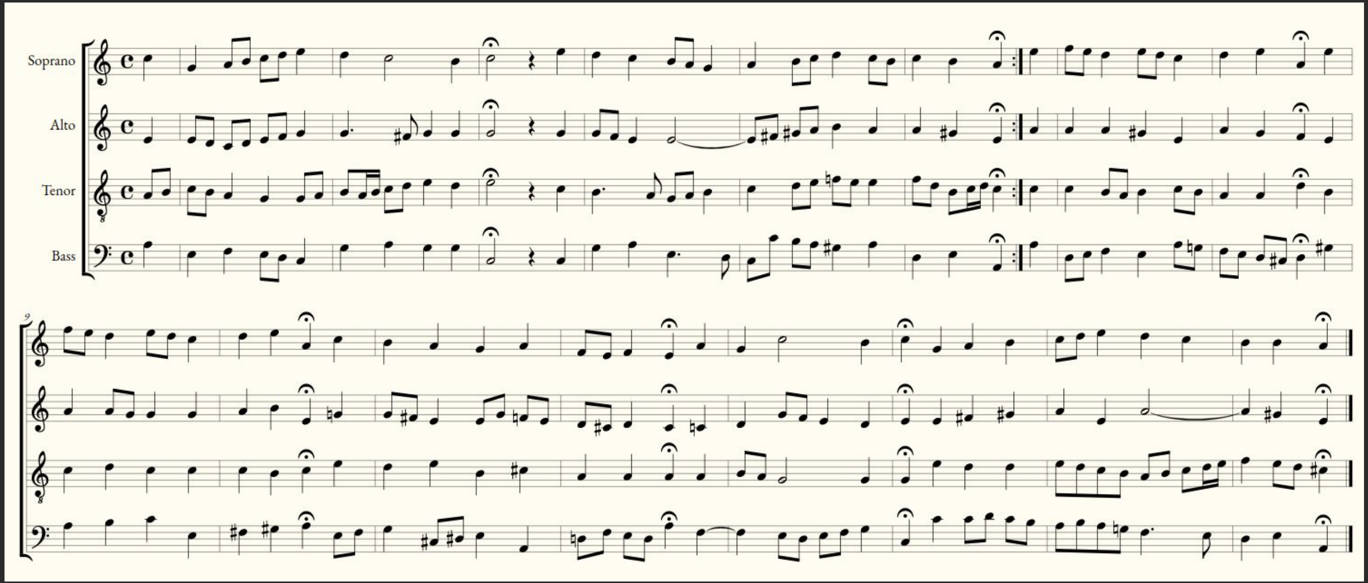
Create Representation System

```
!! midi example #1
**kern
*M2/4
*C:
=1
8c
8r
8d
8e
=2
8f
8g
8a
8b
=3
4cc
4r
====
*-
```

```
!! midi example #1
**MIDI
*Ch1
*M2/4
*C:
=1
72/60/64
36/-60/64
36/62/64
36/-62/64 36/64/64
=2
36/-64/64 36/65/64
36/-65/64 36/67/64
36/-67/64 36/69/64
36/-69/64 36/71/64
=3
36/-71/64 36/72/64
72/-72/64
====
.
*-
```

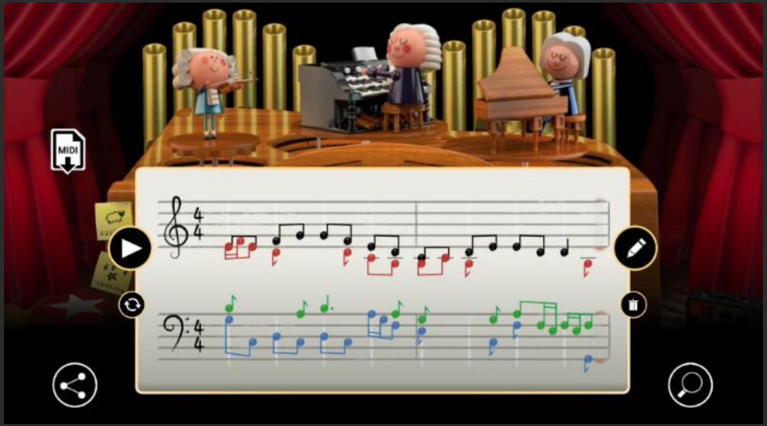
Create Representation System

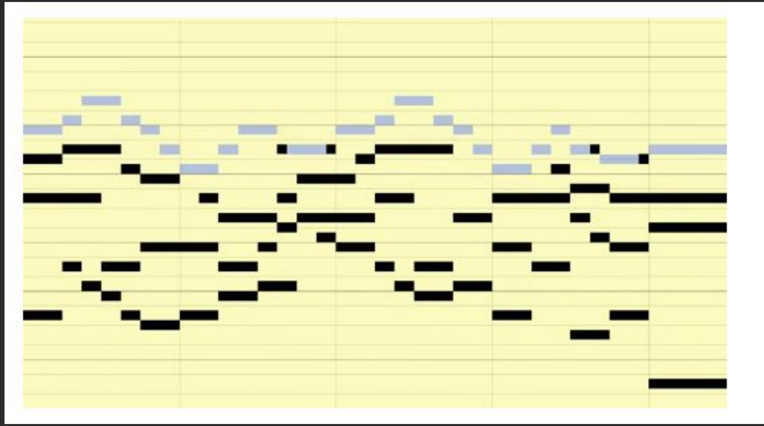
```
!!!COM:     Bach, Johann Sebastian
!!!CDT:     1685/02/21/-1750/07/28/
!!!OTL@@DE: Allein zu dir, Herr Jesu Christ
!!!SCT:     BWV 33/6
!!!PC#:     13
!!!AGN:     chorale
**kern      **kern      **kern      **kern
*ICvox      *ICvox      *ICvox      *ICvox
*Ibass      *Itenor     *Ialto      *Isopr
*I"Bass     *I"Tenor    *I"Alto     *I"Sop
*>[A,A,B]   *>[A,A,B]   *>[A,A,B]   *>[A,A
*>norep[A,B]            *>norep[A,B]
*>A         *>A         *>A         *>A
*clefF4     *clefGv2    *clefG2     *clefG
*k[]        *k[]        *k[]        *k[]
*a:         *a:         *a:         *a:
*M4/4       *M4/4       *M4/4       *M4/4
*met(c)     *met(c)     *met(c)     *met(c
*MM100      *MM100      *MM100      *MM100
4A          8AL         4e          4cc
.           8BJ         .           .
=1          =1          =1          =1
4E          8cL         8eL         4g
.           8BJ         8dJ         .
4F          4A          8cL         8aL
.           .           8dJ         8bJ
8EL         4G          8eL         8ccL
8DJ         .           8fJ         8ddJ
4C          8GL         4g          4ee
.           8AJ         .           .
=2          =2          =2          =2
4G          8BL         4.g         4dd
.           16AL        .           .
.           16BJJ       .           .
4A          8cL         .           2cc
.           8dJ         8f#         .
4G          4e          4g          .
4G          4d          4g          4b
=3          =3          =3          =3
2C;         2e;         2g;         2cc;
4r          4ry         4ry         4r
4C          4c          4g          4ee
=4          =4          =4          =4
4G          4.B         8gL         4dd
.           .           8fJ         .
4A          .           4e          4cc
.           8A          .           .
4.E         8GL         [2e         8bL
.           8AJ         .           8aJ
.           4B          .           4g
8D          .           .           .
=5          =5          =5          =5
8CL         4c          8eL]        4a
8cJ         .           8f#J        .
8BL         8dL         8g#L        8bL
8AJ         8eJ         8aJ         8ccJ
```



https://verovio.humdrum.org/

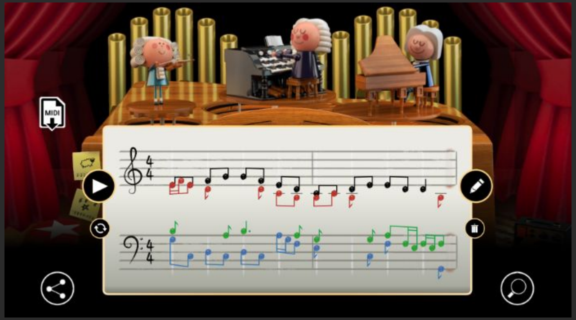
Create Representation System

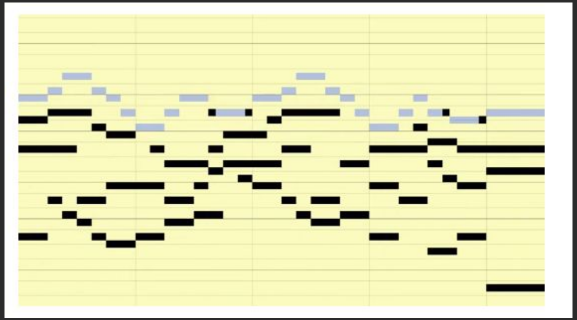


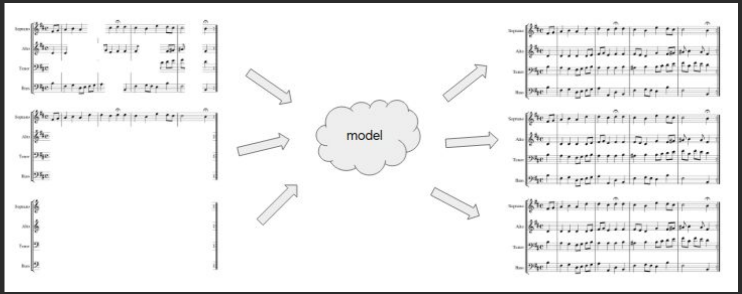


Create Representation System

- Musical “data” needs to get stored in way that computer can read
  - (series of numbers or letters)
- Piano roll representation
  - Pitch on Y axis
  - Duration (on/off) on X Axis
  - MIDI (Musical Instrument Digital Interface)
- Machine Learning Algorithm is trained on “learns” sequences of digital tokens (notes and rests)
- Pattern matching system responds to new input







Wavenet

Representing Music

WaveNet: A Generative Model for Raw Audio

Aäron van den Oord, Sander Dieleman, Heiga Zen[†]

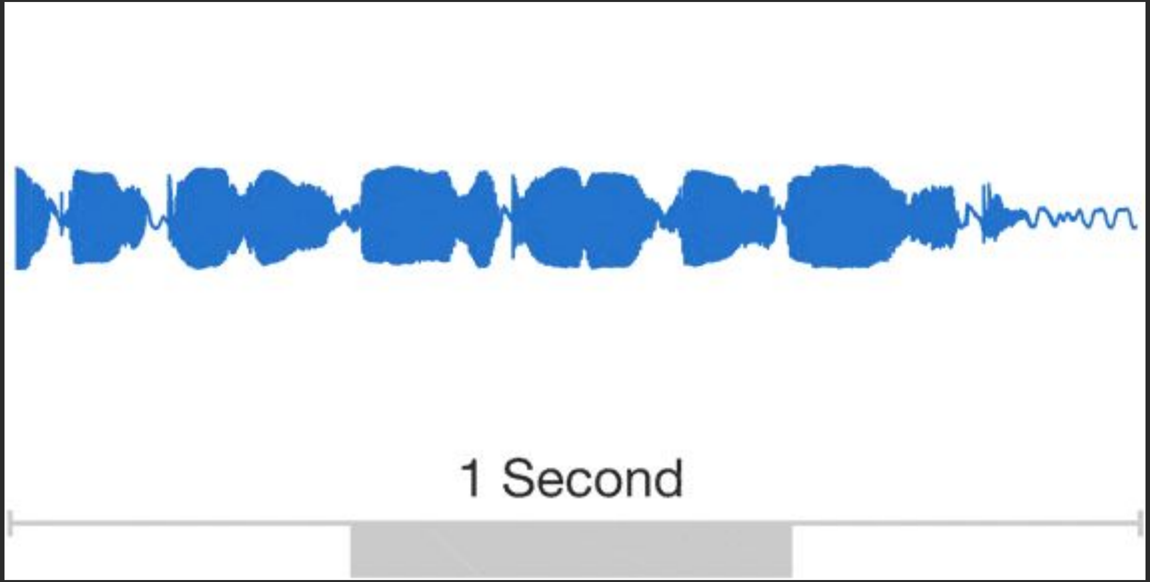
Karen Simonyan, Oriol Vinyals, Alex Graves

Nal Kalchbrenner, Andrew Senior, Koray Kavukcuoglu

{avdnoord, sedielem, heigazen, simonyan, vinyals, gravesa, nalk, andrewsenior, korayk}@google.com
Google DeepMind, London, UK
[†] Google, London, UK

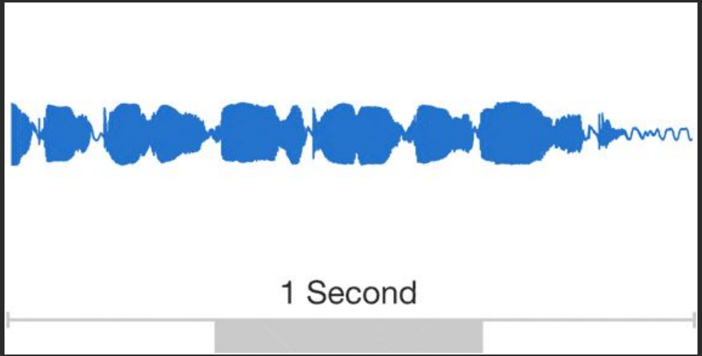
Figure 1: A second of generated speech.

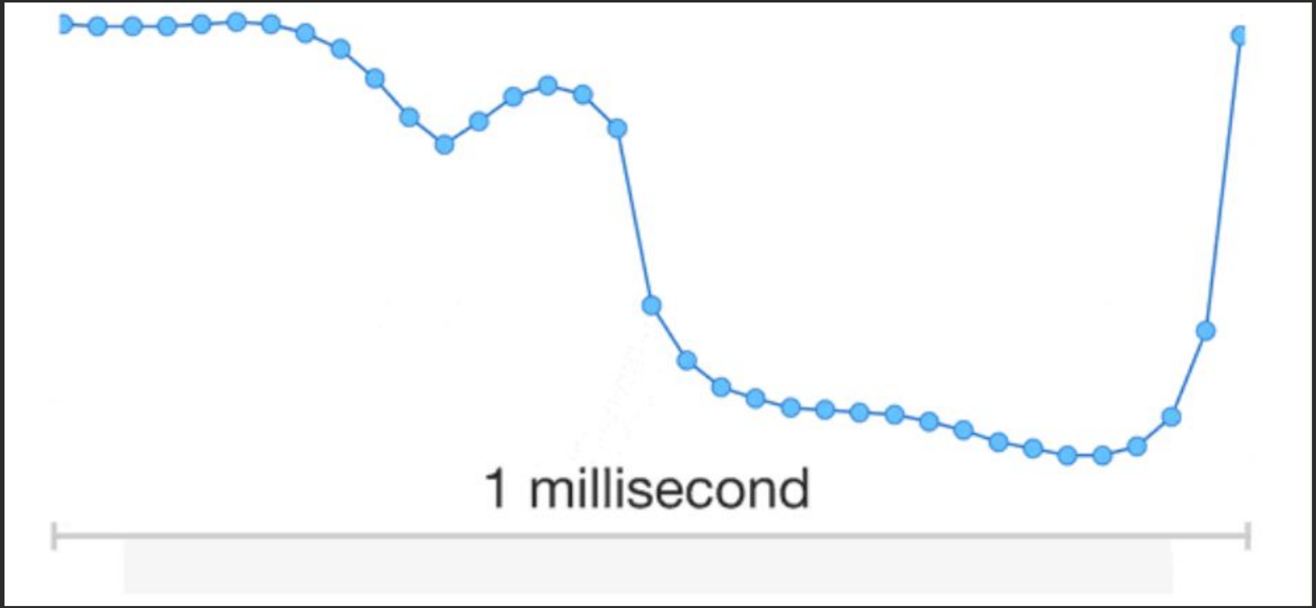
Audio Digitized



Audio Digitized

- Time on X Axis (Like Before!)
- Amplitude on Y Axis





Waveform contains information of patterns that can be learned

Issues with Creating Systems

- Does training a model on a set of musical pieces codify a style?
- What gets included in the corpus that a model is trained on
- Who decides what counts as part of a style and what does not?
- Should the performers of living music have a say in the creation of AI tools?
- Who owns the music that the model is trained on?

Where To Go From Here?

Music and Science Futures

- Field is made up of real people trying to solve real problems
  - Find research that interests you
  - Find people who do what you want to do
  - Contact them!
- Find Societies with common interests, go to their conferences!
  - ESCOM -- European Society of Cognitive Sciences of Music
  - SMPC -- Society for Music Perception and Cognition
  - ICMPC -- International Conference on Music Perception and Cognition
  - ISMIR -- International Society of Music Information Retrieval
- Find online events
  - Future Directions for Music Cognition at The Ohio State University
  - http://org.osu.edu/mascats/virtual-speaker-series/
- Follow people on Twitter #musicscience

Thank you!!