Aaryan Sharma
Assignment 1
21 January 2023

Music, Mind & Technology

Lessons learnt from transitioning from MIR to Cognitive Science

About the speaker

The speaker (Jean-Julien Aucouturier) is formally trained in MIR and is now actively involved in neuroscience/cognitive science. He made several mistakes in his interdisciplinary journey, and therefore he sees the MMMT workshop as the perfect place to talk about his experience. After his first five years of career in MIR, he moved to Japan, marking the transition to a different field of science - Cognitive Science. He then returned to France in 2011 and did minor work on Music Cognition. He was very actively involved in the MIR community, and attended many MIR conferences but has had little to no presence in the field for the last fifteen years.

MIR v/s MCG

Music Information Retrieval (MIR), has a lot to do with signal processing, Machine Learning of sounds and sound synthesis, while Music Cognition (MCG) is mainly concerned with biology, understanding how the brain handles sound, music and voices. Related fields are Psychoacoustics, Cognitive Psychology, Music Cognition, and Cognitive Neuroscience.

These two fields have the same objective: Modelling and simulating how the brain perceives sound or music, but they do that with drastically different methodologies. The speaker is at the interface of both of these disciplines. These two fields also have a very different notion of what constitutes a scientific result, and what is worthy of publication. Sometimes, they even obey a completely different criterion for experimentation.

The talk aim of the talk was to give some insights on how to build scientific proof (that is valid in the field of cognitive science) using the tools of MIR.

Mistakes and Lessons

Mistake #1: Build proof on machine judgements

"Informative about the properties of the algorithm, but not those of human cognition"

This means a paper, which might seem suitable to be published in an MCG journal, might only tell you something about the algorithm you used to come to the result, like the accuracy etc. Papers like these tell nothing about human cognition and therefore are not suited to be published in the field of cognitive science. This is because you do not know whether your algorithm is correct based on the (possibly implicit) models that you have theorised.

Lesson #1: Build proof on information that is available for the machine to process

"Physical information exists in this type of signal/behaviour that's enough to simulate the task"

Motherese is a type of speech associated with an older person speaking to a child or infant. It is usually delivered with a "cooing" pattern of intonation different from that of normal adult speech: high pitch, with many glissando variations that are more pronounced than normal speech.

A long-standing question in cognitive science was, does Motherese help babies to learn a language, and if yes, why and how?

A paper tried to answer this question using tools of MIR like MFCC, SVM, and ML. It proved, using those tools, that there is more information in Motherese in contrast to Adult Directed Speech (ADS), which is what helps the baby brain grasp the human language better. It got accepted in Current Biology, a leading cognitive science journal, as its results were not a function of the algorithm, but of the input information that was available to the system.

Mistake #2: Using signal features to explain human judgements

"Very low explanatory powers: one does not know what is being simulated by the feature"

A paper trying to correlate the signal features of music, like timbre, tempo, tone etc, to the human response will not get accepted as the result of the paper just explains the data and the statistics concludes nothing about the brain or psychology of the human brain. The result is not psychological in nature.

Lesson #2: Explain with intermediary judgements/percepts/action units

"which you model with signal, but allow characterising behaviours in a way that's modular, measurable and verifiable"

The idea is not to explain human judgements but to explain/correlate the "intermediary" cognitive features of the brain, like neuronal activity, which then are known to be the cause of human behaviour. It is papers like these which get accepted in Current Biology.

Mistake #3: Correlation ↔ Causation

"As the study does not provide any mechanistic insights."

"The observed effects are not elucidated in psychological terms"

Keyword: Mechanistic

Key Insight: Need to *manipulate* the system under study

Researchers in the field of cognitive science and the like are asking for causation, not correlation. In essence, what they are trying to tell you is that even though they are correlated, the correlation tells you nothing about the mechanism of why and how they are correlated. Or, in other words, they're saying correlation ≠ causation.

There is a preference for studies that merely draw not just correlations, but also provide a plausible hypothesis behind the inner workings of the phenomena. A paper lacking the latter would be considered of lower priority than one which has both.

This is the secret weapon that once you have in your arsenal, you really start to think like a biologist.

There's even a running joke/meme in the community of cognitive science about this, which is encapsulated in the photo of the tweet below.

"Observing something is not sufficient"

Lesson #3: Build transformation models

"i.e. computational models that can generate/transform the acoustic properties of the voice/music so we can demonstrate (causally) their impact on behaviour/cognition"

Final Reflections

The speaker concluded his speech with his reflections and insights on interdisciplinary research. All of them are summarised into four main bullets below:

- Interdisciplinarity is a very long learning process (cf. the language, then the culture)
- Cherish rejections (they're the only way you'll learn something new) - which means you should even…
- Seek rejections: be stupid (because no one else will)
- Be kind & nurture empathy (it's easy to look down on the other community)