INTERVIEW WITH LARS OHLENDORF

AI DOES ALL BUT NOTHING AT ALL

ARTIFICIAL INTELLIGENCE IN BRAND COMMUNICATION

Interview and text by Elke Wisse

The sound agency WESOUND in Berlin deals i.a. with classic brand sound development. In order to make music selection better for use in brand communication, the company uses the potential of artificial intelligence. In an R&D project funded by the Central Innovation Program for SMEs (ZIM) of the BMWi (Federal Ministry of Economics and Energy), algorithmic music composition under the working title iMozart is made usable in the context of audio branding. Lars Ohlendorf talks about the use of AI, the advantages and disadvantages as well as his experiences.

HOW DOES AI WORK? WHAT CAN AI PERFORM IN THE FIELD OF MUSIC COMPOSING?

The term AI is not defined exactly. But basically, it's about an algorithm finding its own way in different situations and, at best, being able to evolve by itself. The basic function of an AI could be illustrated by a simple compressor circuit:

The input signal is measured and adjusted according to the control circuit - a simple circuit with a steady, responsive behavior. If, in addition, a small mathematics module is added, which constantly calculates a simple regression analysis based on the measured values, then the compressor can already predict foresight. It creates the first patterns and thus the first signs of an intelligent circuit. In a next step, the compressor wouldn't only know the statistics, but also synchronizes with an external database; a memory so to speak. It can thus compare the applied signal and its own control behavior with the patterns stored in the database, and from this derive the best settings for the applied signal – a much smarter approach already. If the circuit then writes independently into this database, so not only uses, but also stores its own data, then we are already touching the areas of machine learning. The program can thus develop by itself.

WHAT DOES THAT MEAN IN THE CONTEXT OF MUSIC?

Music programs that generate music in real-time have been around for a long time. Among other things, these systems are based on the software Autobusk developed by Clarence Barlow. In general, this software does nothing else but to output random MIDI values to trigger sounds under defined limits. For example, I can define how great the melodic ambitus is, which scale the melody should derive from, how strong the scale is represented, and much more. However, for certain associative or activating attributes, for example 'bright' or 'exciting', the program cannot yet generate the music for it because it lacks the corresponding operationalizations. But if the program knew what 'bright', 'exciting', or 'calming' means in musical terms, it could certainly set the musical parameters accordingly. A musical composer would've done the same thing.

HOW DO YOU TEACH THE DATABASE SOMETHING LIKE THIS?

A very recent approach is the one of deep learning. Here you usually work with larger amounts of data and a learning algorithm, a neural network. Music pieces are recorded, and subjects give simple, subjective reactions. At the same time the characteristics of the musical material are examined. The algorithm then searches for even the tiniest correlations between the different measurement data from both questioning and digital signal. The larger the data set, the better the results. For us, however, the data obtained this way would not have any immediate value. The algorithm might perhaps recognize what effect a piece of music has, but it won't be able to produce a piece by itself immediately. Also, one wouldn't know which musical parameters lead to which attributes. It remains a black box. Quite in contrast are the classic, empirical approaches: quantitative and/or qualitative surveys. Particularly promising for us here is the expert survey. Why do composers, as well as design and communication experts, appreciate certain music and how would they compose and orchestrate certain moods? For iMozart we use the already available empirical studies in this area as well as qualitative interviews. Hereby we try to connect musical parameters specifically to brand attributes. Then, we train our algorithm with compositions composed by experienced composers for each individual brand attribute. The algorithm then analyzes the musical parameters to draw the connection to the attributes. In this way, the system can improve itself even after its completion.

WHAT CAN PROGRAMS ALREADY DO IN THIS RESPECT?

All over the world, research is being done on AI in the context of music. It is indeed possible to confine the areas of Music Information Retrieval, Musical Performance and Music Generation. And the results are sometimes very promising. But for decades, a problem outside of AI has been problematic: which parameters in music trigger emotions in humans? I already dealt with this during my studies and quickly learned that this is a field of low significance; it is simply not possible to make flat assumptions. There are way too many variables that cause troubles. After Russell's Circumplex model, basic moods can be classified in the two-dimensional Valence-Arousal model. From this, for example, the following conclusions can be drawn: the faster a music becomes, the more stimulating it is. This can either be happy and jolly or stressful and overstrained. On the other hand, the slower a music becomes, the more soothing, bigger, or even grumpier it acts. Commonly assumed, however, that certain scales had a certain affect associated (e.g. that a major tonality stands for happy and minor for sad emotions) applies only inconsistent. These tonal modes are culturally dependent and geographically influenced. Experiments with infants unknown to major/minor have shown that the major or minor third has no moody significance at all. Much more helpful here are the parameters tempo and rhythm, as suggested in the Valence-Arousal model. These are most likely the ones to allow a prediction of an attributed emotion, because they are direct and experienceable by bodily movement. But it's not like you absolutely know by now what triggers certain emotions in music. That's why it's a challenge to develop a general algorithm that can generate musical stimuli; either to pick us up in certain moods or to put us in certain moods.

DOES THIS APPLY TO ALL MUSICAL GENRES?

Genres are a good keyword! Because here you leave the physical Circumplex area and focus on cultural peculiarities, such as sounds and sonic relationships that supporters of certain cultures have learned and are therefore obtainable. You can build on these. So, we are talking about musical stereotypes – you may say clichés. For example, the sound of a folksy acoustic guitar is often combined with nature and campfire romance. Big band swing with mafia, prohibition and bob hairstyles. Or hip hop with urbanity, battles and underdogs.

This list can be continued endlessly, but it's important to remember that these clichés are of course rough simplifications, which are not always valid for everyone at any given time. Classical music itself is one, big indifferent container, which can be literally anything: uplifted, elitist and arrogant, but just as exclusive, sophisticated and smart. It always depends on the point of view. These are the clichés we try to put into our database to make them available for further use.

IS THERE ANY SOFTWARE THAT ALREADY EXISTS?

There is a whole range of software that can generate music in real time. Composers and researchers around the world are working on the most amazing creations, many of them on platforms like max/msp or puredata. But there are also the first applications for the consumer market. For example, the software Endel is available in the App Store and Google Play. This is an app that evaluates the individual situation of the user through health information from fitness apps, the pulse rate, location data and the time of day to generate matching music. The user can select music to relax, concentrate, fall asleep or to move around. This is exciting but strange at the same time, because you might wonder why you should hear this music. There is nothing that I personally can refer to as a listener, nothing that would appeal to me emotionally. If we are talking about pop music – an AI composition cannot be more than that content-wise – the composition itself is only a small part of the thing called pop music. Pop music is also about the lyrics, the artist's personality, the video art, the artwork, the reception etc. Music created solely by an AI becomes very functional and is thus advertised that way. But that means that our listening skills and ultimately us – as human beings – tend to become rather functional. Music no longer has a reason to exist by itself but should put me in the ideal state of mind for concentrated work, be it in the office, at the weight bench or whatever else you would like to include in the functional universe. Sleep, for example!

IS THERE ANY PROGRAM FOR PROFESSIONAL APPLICATIONS?

Yes, for example, there was a large project at the University Hospital Eppendorf (UKE) in Hamburg together with the University of Music and Theater (HfMT) with Prof. Georg Hajdu and Prof. Dr. med. Eckhard Weymann [1]. Based on the Autobusk algorithm (s. a.), a tool was developed for the background sounds of waiting rooms in the UKE. The question was how to generate music for a long period of time, which is on one hand supportive for therapeutic purposes and on the other hand free of repetition. The Autobusk algorithm has been modified to integrate sensor data such as volume levels and daytime, and then be able to adjust the musical compositions accordingly. Although this is not a real AI application yet, but rather an intelligent responsive system.

WHAT'S BEHIND IMOZART?

iMozart is a composition tool that can generate music in real-time based on a given taxonomy. The software can assign simple attributes such as sporty, rough, or classy to a particular musicality or sonority. The program then creates simple musical structures that define these attributes. We want to use the tool in brand sound projects as communication support between the agency and client. As designers, we have learned to talk about sound for many years, and we know exactly how to represent a brand with musical excitement. But in the creative process, it's about capturing and implementing the client's ideas, not our own, of course. As a rule, you approach a mutual idea of a design by trial and error. If we - and this is the thing - had a tool to playfully make basic musical ideas directly audible in conversation with our client, and if he himself could also change parameters, then we could find common ground much faster. And then

[1] http://georghajdu.de/wp-content/uploads/From-Atmosphere-to-Intervention-final.pdf

afterwards we could communicate a precise, musical briefing with the composer.

It is not about replacing the composer, but about overcoming a communication barrier. To us, iMozart is a design tool to develop ideas with our clients. And it's a very pragmatic approach. We learn a lot from it, on the one hand about our own perceptions, but also how to deal with algorithms creatively. Questioning the AI's results and optimizing the processes makes the work so appealing.

IS WHAT COMES OUT OF IMOZART THE FINAL PRODUCT?

The result is always only the basis for further editing. iMozart delivers simple structures that sound like a lot, because you can influence a lot regarding genre and instrumentation, but the musical content remains manageable. The algorithm does not produce furbished symphonies. And that's not the point of it. Once the structure and the idea of a brand sound has been defined, as previously described, the composer is always required. In this way, the composer gets a very accurate idea of what the client wants. He can then build on top of this with additional creativity and can also deviate from the computer results. This is the only way to create ideas that no algorithm can generate. Because the algorithm will always only deliver what we have fed to it in some way before, or what we already know. Like an echo chamber. We already know this from other sectors.

HOW DO COMPOSERS REACT TO THIS ARTIFICIAL MUSIC? DO THEY SEE IT AS A TOOL OR DO THEY FEEL THAT THEIR CREATIVITY IS RATHER LIMITED BY IT?

So far, I have only received positive feedback. It is a fundamental question of how to develop a briefing with the client, which eventually leads to a certain liability. Listening is - in contrast to seeing - a very individual thing. A drawn square is perceived by everyone as such, but in music, the supposedly exact description of what's being heard relativizes rather quickly. As an agency, we want to make sure that the client gets in the end exactly what he wants and what he expects.

SYSTEMS ALREADY EXIST, THAT ORCHESTRA MUSICIANS ARE ABLE TO WRITE...

Yes, that's right, they do exist, and they create music that is simple or mathematically predictable. Many years ago, my father gave me a software that could write fugues and chorales. You enter a Cantus Firmus and, chop chop, the fugue is finished. Using a deep learning algorithm, for example, it can analyze the ten symphonies by Mahler and generate a piece that is very similar to the originals. This is what was happening with Schubert's Unfinished Symphony, which was finished by an AI. I've never heard this sequel, but frankly, it does not interest me either. What is interesting, above all, is the question of what impact the work in its original form has on the listener. And again, it's about the artist's intentions and its natural limitations. The moment we abolish these things, everything collapses. The creative statement of these AI results is questionable in my eyes. Mahler was a great composer, why would you do something similar now? We should always pursue the goal of creating something new, and therefore there is a need for experienced artists to write original music with meaning rather than just copying clichés. Warner Music has recently released an album produced by AI Software Endel. This doesn't make any sense to me either. After all, generative music with AI is exciting because it keeps writing itself and never repeats. If the music is released on a CD for example, it loses its special feature. Mind you, the only superb feature it has.

WOULD THE NORMAL LISTENER HEAR THE DIFFERENCE BETWEEN AI AND HAND-MADE MUSIC?

I don't know if I would hear the difference myself. If, for example, an AI composition is played by an actual orchestra, something artistically valuable can happen, if only because experienced musicians are playing it. This is similar in pop music. Without vocals you probably wouldn't hear the difference, because pop music is in some parts so heavily formatted, that it is algorithmically relatively easy to grasp.

WHERE IS COMPOSING WITH AI HEADING TO IN THE NEXT COUPLE OF YEARS?

If things go well, we will get more and more tools to help us develop musically. I see iMozart as a project, that takes a role in this process as well. If we use the AI creations as a basis and continue working with actual composers, then we can only expect better things. If we see it as a tool, then AI is a very good support. Of course, it could also happen that the result from the AI is simply used like that as the final product. In that case, the artistic quality of composition would go all the way down in the long run. In the end, it will probably be a combination of both. As designers and listeners alike, we are all responsible to steer this development in the right direction.

ADDITIONAL TEXTS:

Lars Ohlendorf is a sound engineer, sound designer and composer with twenty years of professional experience. He specializes in 3D audio, interactive design and generative systems. At WESOUND, Lars is Creative Director / Head of Design and is responsible for the development of sound concepts and music strategies. Lars first visited the SAE, then took the master's degree in Sound / Vision at the HAW in Hamburg.

Indication of source

WESOUND GmbH (2019)
www.wesound.de

German original in: VDT-Magazin, Issue 3/2019 (July 2019), pp. 32-36.
www.vdt-magazin.de