Internal Report 2010-10 August 2010

Universiteit Leiden

Opleiding Informatica

Art or Artificial

Harma Everts

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Art or Artificial

Harma Everts

August 2010

Bachelor Research Project Computer Science

Supervisor: Edwin van der Heide

Universiteit Leiden

LIACS

Abstract

Are people able to distinguish computer generated random dodecaphony from human composed dodecaphony? A computer program was developed to generate random dodecaphony, according to certain templates, distilled from human composed music. People from different musical levels, ranging from laymen to experts, were then asked to listen to both human composed and computer generated fragments, and to indicate what they thought they were listening to, composed or generated music.

The obtained results showed an average score of 50.46% correct answers. Since a correctness of 50% could also be reached by guessing blindly, it seems that on average people cannot hear the difference between generated and composed dodecaphony. The musical level of the respondents however did make a small difference. People with more listening and performing musical experience on average correctly identified more fragments, but usually with a small majority, as opposed to the people with less musical experience, who on average achieved the highest score.

This could be explained by the importance of the twelve tone row, and the attention to detail of the more musically skilled respondents. They would be expected to notice more on a detailed level, whereas the laymen would intuitively give an answer according to their overall feeling of the fragment. Over thinking and looking for sequences and thematic material may have led some professionals to believe that some generated fragments were really composed by a human being.

More information is needed on the attributes that define the human factor in a composition. This experiment should also be repeated on a larger scale, to confirm or invalidate the results obtained in this small research project. The most important conclusion of this research project therefore must be the fact that more research is needed to explain the obtained results.

Contents

1 Introduction

In this thesis, we will describe the research project "*Art or Artificial*". The project researches whether people are able to distinguish computer generated random dodecaphony from human composed dodecaphony. A short introduction to dodecaphony can be found in section 2.

The computer program Compunist was written to perform the research. This program is able to randomly generate dodecaphonic pieces according to a scope and parameters which will be explained in sections 4 and 5.

Section 6 describes how the experiment took place, and the results can be found in section 7. Some conclusions will be drawn in section 8, and section 9 describes some possible future research.

2 Dodecaphony, a short introduction

Dodecaphony, or twelve tone composition, is a composition technique invented by Arnold Schönberg in 1921. The limits of common harmony were reached, and there was a need for a new way of composing and creating consistency. Schönberg and his colleagues Anton von Webern and Alban Berg formed the Second Viennese School and laid a basis for serialism and contemporary classical music.

The basis of a twelve tone composition is an ordered arrangement of the twelve tones of the chromatic scale. There are four rules that apply:

1. All twelve tones of the chromatic scale must occur once.
2. No note is repeated in the set.
3. The set may be stated in any of four transformations: prime, inversion, retrograde and retrograde-inversion.
4. The set, in any transformation, may be modulated.

The tone row chosen as the basis of the piece is called the prime series (P). Untransposed, it is notated as P_0. Given the twelve pitch classes of the chromatic scale, there are (12!) (factorial, i.e. $479.001.600^{10}$) tone rows, although 469.015.680 of these are merely transformations of other rows. There are 9.985.920 truly unique twelve-tone rows.

An example can be seen in Figure 1-4.

Figure 1: The Prime row

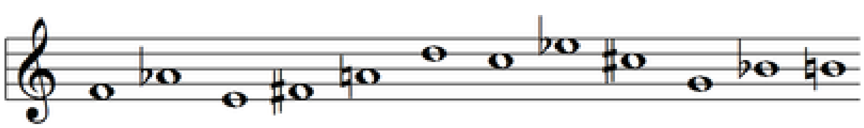


Figure 2: The Retrograde (Prime in reverse order)

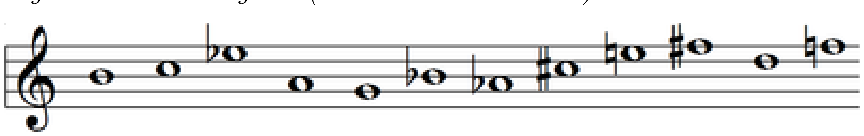


Figure 3: The Inversion (Prime upside down)

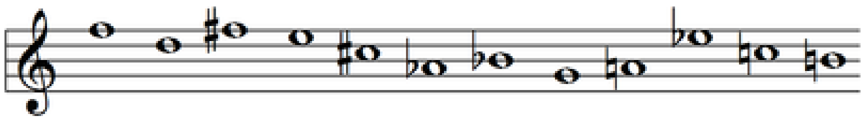


Figure 4: The Retrograde-Inversion (Inversion in reverse order)

3 Expectations

Since the nature of dodecaphony is quite a clear algorithm, it might be easy to mimic by a computer program. A human composer however may create a tone row that has a clear center point, or has quite a recognizable theme. Will the difference be discernable to the listener?

Dodecaphony is atonal in nature and it is difficult music for the average listener. Most people don't have much, if any, experience listening to it. It is hard to hear structures in atonal music when all you've heard before is Bach, Mozart or pop music. The basis of dodecaphony, however, is very strong: the twelve tone row. Even if not heard on a conscious level, it should provide patterns and structures within the piece, especially when combined with a very strong rhythmic theme.

When creating random dodecaphony, themes will only appear randomly and are not generated on purpose. The rhythm will be generated randomly as well. This may result in the pieces sounding erratic and confused, even though there also may be the occasional well-sounding moment.

In a human composed piece, it is expected that the composers thoughts about everything, like choosing the twelve tone row, creating rhythmic counterpoint and creativity in the structure, will be audible. The computer will generate an (on average) uninspired piece of music, since there is no real building up of tension, no real variation, no parts, and no real theme. It is however doubtful whether the laymen will be able to hear this, but the more advanced musicians should certainly be able to hear the difference.

What makes a (dodecaphonic) piece sound human composed? This is a difficult question to answer in this case, and we will describe some possibilities:

- The basis, the twelve tone row, has been chosen carefully. It provides either a building up of tension, some very clearly important notes in its structure, a hint of tonality, and in general a basis for a very interesting piece of music. The composer chooses the tone row based on a succession of notes which builds up and relaxes, in all transformations, and the different transformations should add extra dimensions to the musical material.

- The rhythms are balanced. A piece may sound very free, or very rhythmical, or a combination of both with a clear contrast. There isn't much clutter and when there's a sequence, the rhythm is usually part of it and the same with every iteration of the twelve tone row, to enhance its recognizability.

- The using of rests is carefully thought out. When the piece sounds very polyphonic, the rests are usually more scattered around and not general pauses. Otherwise, the voices have rests together most of the time, to bring out certain structures in the piece more clearly (different parts, different atmospheres).

- There either is careful usage of harmonies, or of dissonants, or both. When using a lot of dissonants, the tension is brought to a maximum. Harmony or the hint thereof may be used to relax those tensions.

- The piece can sound very different in different parts. A piece generated by Compunist will sound more or less the same throughout, since the parameters and chances stay the same throughout the whole piece.

All these things put together create a musical piece which has a clear overall structure and balance. It is clear that every note in the row, every rhythm and chord has been chosen with a purpose. This, ideally, is how a human composed piece sounds. In short, there is a lot of structure in all musical parameters. A randomly generated piece only occasionally satisfies these requirements, and a truly well-balanced piece only occurs as a coincidence, not as a rule.

Dodecaphony may sound erratic to a layman's ear, and maybe even to the ear of a well-trained professional. Random dodecaphony may sound just as erratic, but there might be a difference with the composed music. This difference may be important for defining the line between human and computer. It is also possible that there is a difference, but that the average listener cannot hear it.

4 Creating random dodecaphony with Compunist

Compunist is a C++ program that has been developed especially for this research project. It is capable of producing random dodecaphonic music, based on pre-defined style parameters. The piece that is created is quite random in nature. The output of Compunist is a Lilypond [1] file. Lilypond is a typesetting program for musical scores, and it can produce postscript, pdf and midi files. Thus the compusition can be transformed into either an audio file or sheet music or both.

Compunist creates a twelve tone piece, based on the parameters which will be discussed in section 5 of this document. The basic twelve tone row will be chosen randomly, and the generation of the piece will be based on randomness and chance. In the experiment, the produced pieces will then be played as midi files, together with some human composed pieces. It is interesting to see whether both laymen and musicians will be able to make a distinction between the two.

Compunist offers two ways of generating a piece: the user can select a template based on an existing composition, where all parameters are prefilled, or the user can create a custom compusition from scratch by giving in their own parameter values. This is done in a menu-like structure.

4.1 Scope

It is the intention of this research project to find out whether twelve tone music generated by a computer can be discerned from twelve tone music composed by a human being. To be able to compare the two, certain choices had to be made with regard to the scope of this project.

4.1.1 Midi files

The musical pieces would all be presented to the test panel as midi files. Since the generated music is quite hard to play, even for skilled musicians, it would take too much time and effort to prepare a set of live-played test files. Also, by playing the pieces as a human, the human component re-enters. A human being playing music always interprets it in a certain way. This could bias the research, and make the results unreliable. We chose to keep the audio fragments clean, even sterile. The midi files all have the same generic piano sound.

4.1.2 No dynamics

To recognize the style or structure of a piece, the notes and the rhythm are the most important aspects. Playing a piece with wrong notes or a different rhythm will make it unrecognizable. On the other hand, leaving out the dynamics will keep the structure of the piece intact. Playing *Eine kleine Nachtmusik* without dynamics will not cause people to fail recognizing it, but changing the characteristic rhythms or melody would.

4.1.3 Ignoring the instrument specifics

In the implementation of Compunist, the different voices are what they are, voices, with a certain range each. However, the playability (fingering) on either a piano or any other instrument is not implemented. This seemed unnecessary, since the test files would be midis after all. Also, the different voices could always be sung, since the different ranges correspond to the human voice ranges they were named after.

4.1.4 No intended harmony or polyphony

Compunist creates random twelve tone melodies for each voice. The voices, however, don't interact, and are created individually without looking at the other voices. This means, that any harmonies or polyphony occurring will be completely random. In this aspect, the generated pieces are quite two-dimensional.

4.1.5 Using templates

To keep the pieces from being totally random, certain parameters from the existing compositions were used as a template to create random pieces. Those parameters were determined by carefully analyzing the pieces of Schönberg, Webern and Stravinsky, and will be clarified in section 5.

5 Specification of the Compunist parameters

Certain parameters had to be specified, to be able to create twelve tone pieces that follow the rules, but still are as random as possible. The following list contains the parameters used:

- Twelve tone row (created randomly)
- Note lengths
- Rest lengths
- Tempo
- Number of voices
- Length of the piece

One can think of more advanced parameters, like dynamics, themes, sequences and changing time signatures, but those are not essential to create a basic random piece of music. The parameters stated above can be set by the user when they decide to create a custom template for a twelve tone piece. In the pre-defined templates, all parameters are set automatically.

The different values used in the templates are based on the different pieces as a whole. If in one of the pieces one of the voices has mostly rests and the other has mostly notes, the average percentage of notes and rests is taken for the whole piece. Compunist doesn't try to mimic the original pieces, but uses parameters which in theory could produce the pieces they were based on.

Throughout the program, a random generator is used to make certain decisions based on chance. The random generator is initialized with a time stamp, ensuring that the compusition will be different every time the program is used.

5.1 Twelve tone row

The twelve tone row is randomly generated. The twelve notes are represented by the numbers 1 through 12, and will be selected one by one by a random generator. At the beginning stage, it is also decided which accidentals to use (flats or sharps). The Inverse, the Retrograde and the Inverse Retrograde are derived from the Prime row. Which of the four transformations to use is decided randomly.

There are three ways of dealing with the twelve tone row. The first, strict one prescribes that every note may be used only once in every iteration. The second one is a little less strict, and allows the repetition of the current tone. The third one is the less strict, even the previous note may be repeated any number of times. Even Schönberg himself loosened the rules after his first few twelve tone compositions and most composers who make twelve tone music at some point used the rule that the current and/or the previous tone may be repeated.

5.2 Rhythm

The following notes and rests can be used:

- Whole note/rest
- Dotted whole note/rest
- Half note/rest
- Dotted half note/rest
- Crotchet/rest
- Dotted crotchet/rest
- Quaver/rest
- Dotted quaver/rest
- Semi quaver/rest

For each note and rest, a usage percentage will be provided. Together, these will be a hundred percent. The percentage represents the chance that the next note or rest will be that particular note or rest. Tuplets will not be used.

5.2.1 Denseness

A musical piece consists of both notes and rests. To create a certain atmosphere, the balance between the two can be adjusted, by not dividing equally between the chances of notes and rests.

5.2.2 Tempo

The atmosphere of the piece is heavily influenced by the tempo in which it is played. The tempo is defined as the number of beats per minute. One beat can be represented by any note value.

5.3 Number of voices

The number of voices that Compunist can create is set to a maximum of ten. Each of these voices has a certain range. To keep things simple, one can choose from the six voices S(oprano), M(ezzo), A(lto), T(enor), B(ariTone) and B(ass). Their respective ranges are:

- S: d' to a"
- M: b to f"
- A: g to d"
- T: d to a'
- BT: B to f'
- B: G to d'

The voices chosen can be any combination of the six ranges, and every range can be used multiple times.

5.4 Handling the tone row

The program will create the different voices separately. The voices will not each have their own tone row, but will share the same basic one, although they may use any transformation of it independently. It is not possible for one voice to start two notes simultaneously, and the order of the row must be preserved. Compunist will add a note or rest one at a time to each voice, until the desired length of the piece has been reached. The compusition will be saved during this process, and finally Lilypond output will be created.

5.5 Length of the piece

The length of the piece will be accounted by counting the number of row repeats per voice. The voices might not end at the same moment, since the lengths of the notes are random. When creating a long enough piece, the endings of the voices should still be close to one another, since the chances are equal for each voice.

5.6 Time signature

The piece will have a time signature. Due to time constraints, it was too much to properly implement the Lilypond PDF output, but a score can easily be created by opening the midi in a music notation program like Finale [3] or Sibelius [4].

5.7 Dynamics

No dynamics have been implemented. The bases of a piece are its notes, rests and rhythm. Stripping the human composed pieces from anything human except for that basis should still provide enough information to hear whether the piece was composed or generated. For instance, "Twinkle, twinkle, little star" without dynamics is still recognizable, but playing that song with a totally different rhythm, or changing the notes, would take away the essence of the song.

5.8 Themes

Even though it is quite possible, no real theme repetition or the like are implemented. The goal is not to create a piece that sounds human, the goal is to create a piece that sounds as random as possible, and compare it to human composed pieces. A theme can occur, though, when two voices start sequentially with the same mutation of the twelve tone row.

5.9 Templates

To generate pieces that are in basis quite close to the selected human composed pieces, three templates were created, and their parameters were carefully tested and selected.

The first template is based on the ”Gavotte” from Schönbergs opus 25, a piano piece in five parts. Here, the twelve tone rules are still quite strict, since it is his first real twelve tone piece. The previous note may not be repeated.

The second template is based on the ”Kinderstück” by Anton Webern. This piece is rather simple, playful and rhythmical. The current tone may be repeated, but not the previous one.

The last and third template is based on ”The Owl and the Pussycat” by Igor Stravinsky. Here, the previous tone may be repeated (even several times), which creates a more tonal feel to the piece. The decision was made, not to go into creating a vocal piece per se, since a vocal piece is based on text, and the program is not implemented to understand text. The composition will be treated as an instrumental piece instead.

5.9.1 Schönberg

These are the parameters of the Schönberg template:

- 4 voices: SATB
- Length: 3 repeats
- Flats
- Tempo: crotchet = 44
- $\frac{2}{2}$ time measure
- No repetition of the current or previous note

Length	Note	Rest
Dotted whole	0%	0%
Whole	0%	0%
Dotted half	0%	0%
Half	2%	2%
Dotted crotchet	5%	0%
Crotchet	15%	2%
Dotted quaver	8%	0%
Quaver	50%	4%
Semi quaver	10%	2%

Table 1: Chances of different notes and rests in the Schönberg template

5.9.2 Webern

These are the parameters of the Webern template:

- 3 voices: SAB
- Length: 4 repeats
- Sharps
- Tempo: crotchet = 50
- $\frac{3}{4}$ time measure
- Repetition of the current note is possible, but not of the previous note

Length	Note	Rest
Dotted whole	0%	0%
Whole	0%	0%
Dotted half	0%	0%
Half	1%	12%
Dotted crotchet	2%	0%
Crotchet	4%	22%
Dotted quaver	0%	0%
Quaver	24%	11%
Semi quaver	24%	0%

Table 2: Chances of different notes and rests in the Webern template

5.9.3 Stravinsky

These are the parameters of the Stravinsky template:

- 2 voices: MB
- Length: 3 repeats
- Sharps
- Tempo: crotchet = 60
- $\frac{4}{4}$ time measure
- The current and the previous note may be repeated

Length	Note	Rest
Dotted whole	0%	0%
Whole	0%	0%
Dotted half	0%	0%
Half	4%	0%
Dotted crotchet	0%	0%
Crotchet	10%	3%
Dotted Quaver	6%	0%
Quaver	75%	2%
Semi quaver	0%	0%

Table 3: Chances of different notes and rests in the Stravinsky template

In the Stravinsky template only two voices were generated. In the original piece, there are three voices, but the lower two voices just play in octaves and don't have a separate existence. Thus, there are only two voices with unique tonal material, and the choice was made to omit the third voice. It could be added easily by adding octaves to the lower generated voice.

6 Experiment

To test whether people can hear the difference between the generated pieces and the composed pieces, an online questionnaire was created with ThesisTools [2]. A copy of the questionnaire presented to the respondents can be found in the Appendix of this document. In the questionnaire, people were asked to provide amongst others their musical level, gender, age and their familiarity with dodecaphony. They got nine short midi fragments presented to them. All fragments were played by the computer, so no difference could be heard in the performance of the different fragments.

Three of the fragments were actual compositions, whilst the other six were randomly generated by Compunist. Of these six, two were generated with Schönberg parameters, two with Webern parameters, and two with Stravinsky parameters. The test subjects did not know how many of the fragments were generated and how many were composed by human composers. It would have been possible to pair up the pieces per composer (template), play all three fragments, and then ask the respondents which one(s) they thought were human. That way, the respondents would have had more of a context, but it would take away the spontaneity of just getting one fragment and deciding about its origin immediately.

The testers were asked to fill out how sure they were about their answers, and what the basis was for their judgment. This last question was not obligatory, however, most people did provide an answer.

For the experiment, only parts of the generated and composed pieces were used. The first twenty seconds were played, so there is no real ending to the fragments. Since every piece was originally longer than twenty seconds, all fragments suffer from this lack of ending. No fading out was used, just a clean cut.

Of the tests, 48 out of the 60 entries were usable, the rest of the entries were either only partially filled out or completely empty. The 48 test subjects were quite evenly divided over the different levels of musical skill. This made comparing the different levels easier.

7 Results

7.1 Attributes of the respondents

After the experiment, there were 48 usable results from test subjects. Their attributes can be viewed in figures 5 through 7.

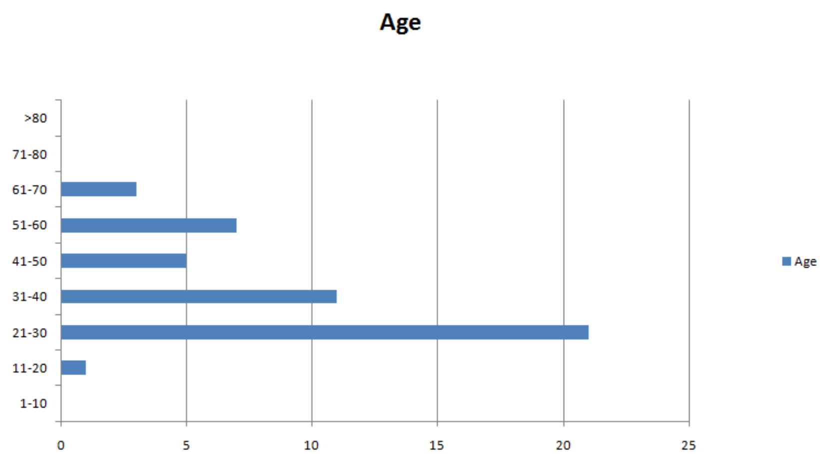


Figure 5: The age of the respondents

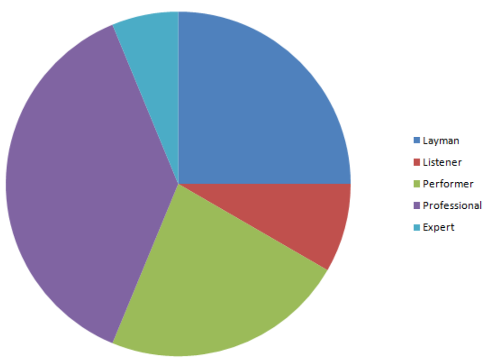


Figure 6: Musical level of the respondents

- A layman can't read notes and never listens to classical music.
- A listener listens to classical music from time to time, but doesn't play an instrument.
- A performer plays an instrument on an amateur level and might listen to classical music from time to time.
- A professional studies at a conservatoire, or has a degree in music.
- An expert is a classical composer or conductor with quite some musical experience, possibly also in the field of dodecaphony.

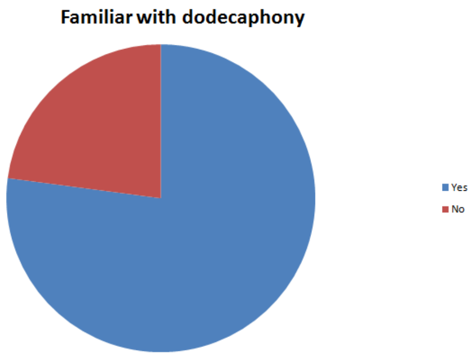


Figure 7: Familiarity of the respondents with dodecaphony

7.2 Overview of the results

In figure 8 the overall percentage of correct and incorrect answers with each audio fragment can be seen. The different fragments were either the real Schoenberg, Webern and Stravinsky pieces, or one of the two generated pieces per template.

1. Webern
2. Compunist (Schönberg template)
3. Stravinsky
4. Compunist (Webern template)
5. Compunist (Stravinsky template)
6. Compunist (Schönberg template)
7. Schönberg
8. Compunist (Stravinsky template)
9. Compunist (Webern template)

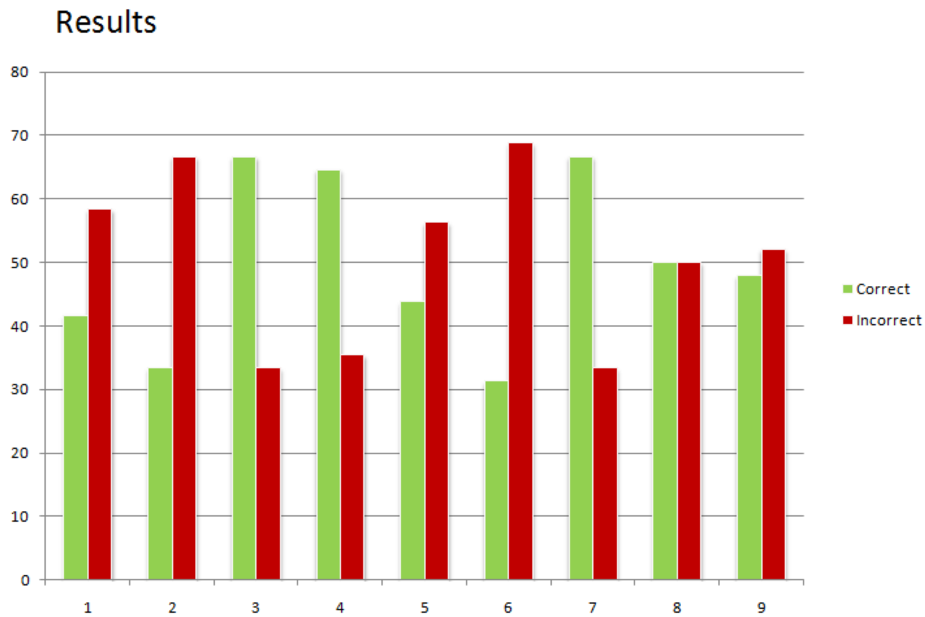


Figure 8: Overall test results per fragment

Two of the real pieces were mostly guessed correctly. Two thirds of the respondents recognized these fragments as being human composed. The Webern piece was not recognized correctly by almost 60% of the respondents, and one of the generated Webern template pieces was often mistaken for being a human composed piece. The rest of the pieces were mostly not recognized for what they were. The eighth fragment (Stravinsky template) was a tie, half of the respondents thought it to be a generated piece, and the other half thought it was a human composed piece.

If we take a look at the results combining each composer with the templates created with that composers parameters, the results are as follows:

Composer	*Identified correctly*	*Not identified correctly*	*Tied*
Schönberg	1 (Composed)	2 (Generated)	0
Webern	1 (Generated)	2 (1 Generated / 1 Composed)	0
Stravinsky	1 (Composed)	1 (Generated)	1 (Generated)

Table 4: Overview of the correct and incorrect answers per composer (template)

The respondents thought that both the real Schönberg and the two generated Schönberg pieces were human composed. One of the Webern generated pieces was identified correctly, the other and the real Webern were not. The real Stravinsky piece was recognized as being human composed, and there was a tie on one of the generated Stravinsky pieces.

Overall, the average accuracy of the answers was 50,46% correct. This is not significantly higher than when every answer is guessed. The Stravinsky piece was correctly identified most often. This may be due to the somewhat more tonal nature of that piece.

When looking at the results per musical level, we get the following results, as can be seen in figure 10:

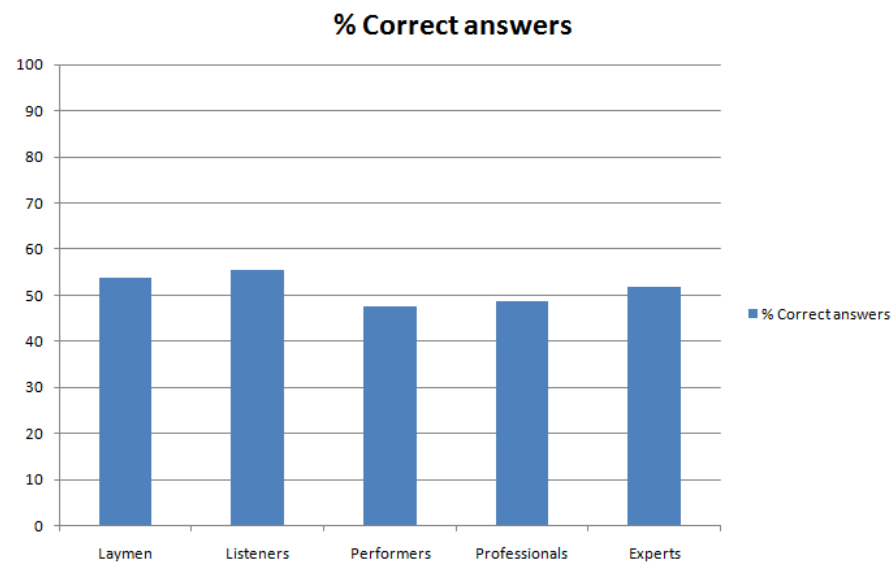


Figure 10: Average percentage of correct answers per musical skill level

It is striking to see that the lower skilled respondents on average have a (marginally) higher score than the higher skilled respondents. It would be interesting to take a closer look at those results. The combined graphs of the two lower levels, the graph of the medium level, and the combined graphs of the two higher levels can be seen in figures 11-13.

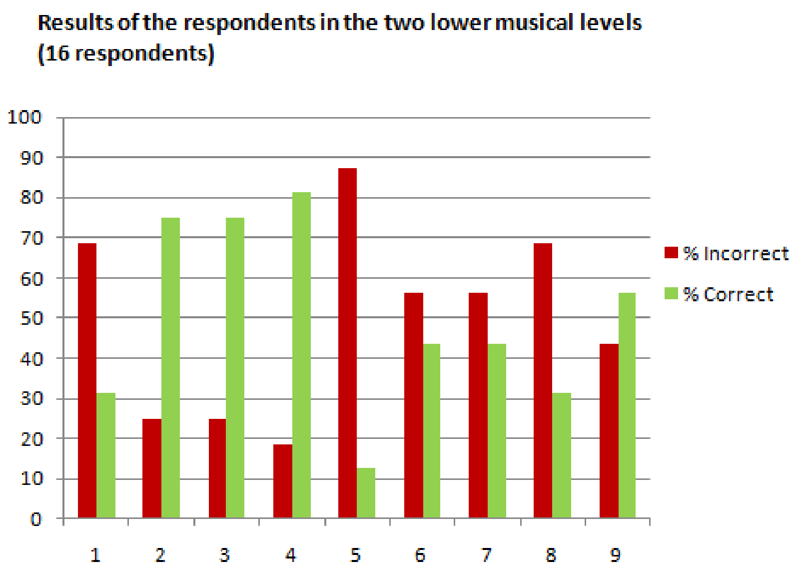


Figure 11: Correctness of the answers per fragment in the lower two levels combined

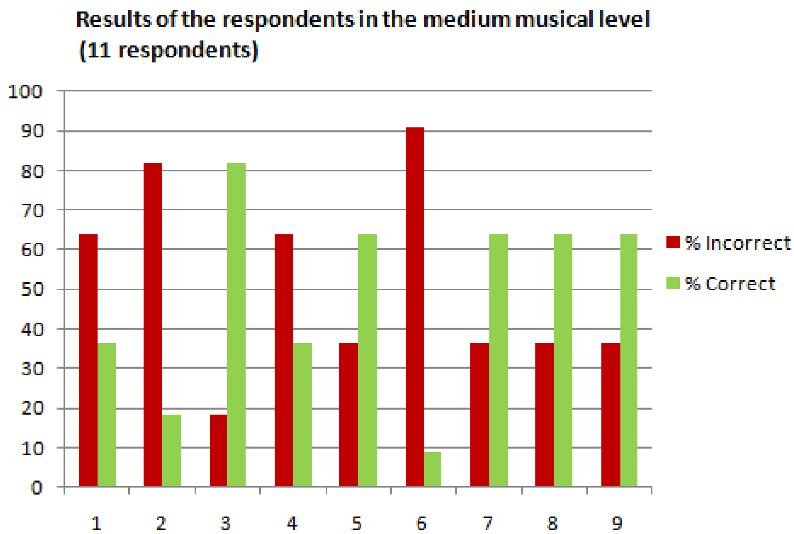


Figure 12: Correctness of the answers per fragment in the medium level

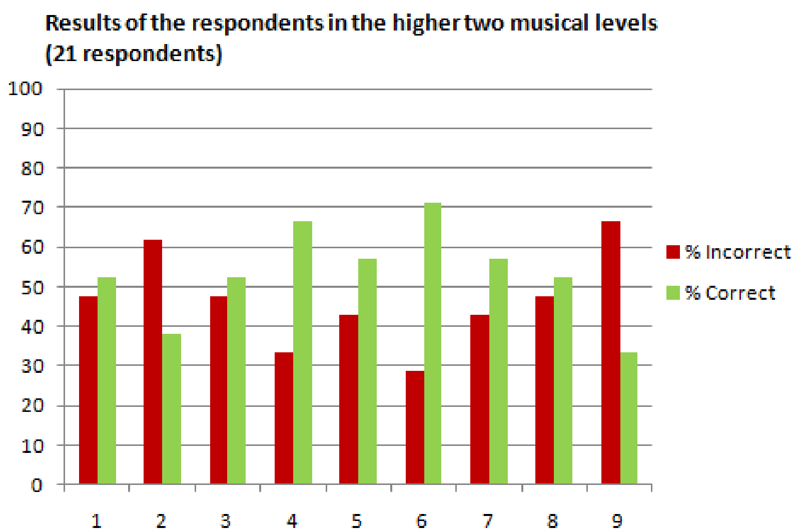


Figure 13: Correctness of the answers per fragment in the higher two levels combined

When looking at these figures closely, it is clear that the graph of the lower two levels is much more erratic than that of the medium level, and that the higher two levels produce the most even divisions between the answers. The higher two levels identified seven out of nine fragments correctly, but mostly with only a slight difference. The lower levels identified the Stravinsky piece correctly, but failed on both the Webern and the Schönberg piece. The medium level identified the Stravinsky piece very well, and the Schönberg piece was also recognized as being composed by a human being.

When listening to the fragments 2, 5 and 6, there are some attributes that may have fooled the respondents. Fragments 2 and 6 were both made with the Schönberg template, and they both have their occasional tonal harmony. They also both seem to build up to a goal. Fragment 2 was only correctly indentified (with a big majority!) by the lower level respondents. Fragment 6 mostly fooled the lower and medium levels, while the higher level respondents achieved their best result there. Fragment 5 only fooled the lower levels, with its strangely tonal start, building up to nowhere, while the medium and higher levels saw right through it. The hint of tonality may be the key here, when it's overdone, the lower levels recognize it, and when it's ”subtle”, the higher levels mistake it for being human composed.

On average, the lower levels identified four fragments correctly, the medium level recognized five, and the higher levels gave seven correct answers. However, the higher levels only provide a slight majority in most cases, so the significance of these correct answers is lower than that of the lower levels. Thus the overall average results of the higher levels are still a little worse than those of the lower levels. Not every level has an equal number of respondents, which also helps explain the different outcome of the overall correctness (figure 10).

8 Conclusions

Apparently, people in general cannot hear the difference between randomly generated and human composed dodecaphony very well. However, there are some sidenotes to make, and there are also some possible explanations for these results. The main conclusion, however, is, that there are some subjects that need more research.

When looking at the composed pieces, two out of the three were mostly identified correctly. The one that often was not recognized as human composed, was the first fragment of the questionnaire. The fact that it was the first fragment people heard may have influenced the score, since the respondents had no context yet, and the fragments being midi files may have influenced the perception of the first fragment more than of the other fragments. When leaving out the first fragment, the overall average correctness of the answers doesn't change significantly.

The rest of the fragments, which were computer generated, were often fooling the respondents. Fragments 2, 5 and 6 were often mistakenly identified as human composed. Those fragments clearly didn't sound that different from the human composed fragments. Of course, taking away the dynamics, playing only a fragment of the piece (with no real end) and the lack of human interpretation took something away from the human composed pieces. The structure however is still mostly there, and the computer generated fragments lack the human finishing and thought. More importantly, the computer generated fragments were composed in a linear way, whilst in the human composed pieces the harmony of the different voices is also part of the composition.

Compunist is not that advanced a program. It only generates random tone rows with random lengths and random rests, and doesn't take the context of the different voices into account. It doesn't try to construct thematical content, or sequences, or any buildup whatsoever. These generated files should not be able to be confused with human composed music this easily.

One could research whether the more musically skilled respondents have had their experience turn against them, listening to the pieces in detail, and trying to detect structures which were there, indicating whether the piece was composed or generated. The less experienced respondents may have listened to the pieces as a whole and felt the lack of consistency in the generated pieces, despite the sometimes thematic buildups. An indication of this can be seen in their comments; the more musically skilled they are, the more the respondents indicate that they tried to find themes and structures, and the lower the level, the more they wrote about the atmosphere of the pieces. The difference between the results of the respondents on average however was not that significant.

There may be several reasons for the fact that the generated fragments were mistaken for real compositions.

1. In some of the fragments there may have been an indication of a "red line" throughout the piece, when several voices start one by one and use the same transformation of the tone row with roughly the same rhythm.

2. From time to time, the generated voices may coincidentally come together in a slightly harmonic way which may fool the listener.

3. The strongest attribute of dodecaphony is the tonal row, which is also strongly present in the generated pieces. The differences between the generated and the composed midis therefore become much smaller, and less detectable.

These explanations may or may not be true, and more research is needed.

9 Future research

To really understand the results of this research, more research is needed. A clear definition is needed of how someone can hear whether a piece is human composed or not. If Compunist can fool so many people already with random output only, what will the results be when the algorithms used in Compunist become more complex? The rules from section 3 can be implemented mostly without a problem. We have to find out where the line is, if any, between computer generated and human composed, and what its definition is.

It would be interesting to do this same experiment on a larger scale, taking for example ten human composed pieces and twenty generated pieces, and having several hundreds of respondents. Would the results still roughly be the same, or are the results of this research project not generic enough due to the smaller scale? One would also need to take the musical skill levels of the respondents into account again, to see whether the results seen in this research project can be confirmed.

Maybe it would make a difference if every respondent would hear the different fragments in a random order. In that case, it would be easy to see, whether the position in the experiment makes a difference to the judgment of the respondent (especially when the fragment is the first fragment). It might also make a difference if the full pieces were played, instead of only part of the pieces. That way, the overall structure of the piece might become clearer, and the respondents might be able to identify the fragments correctly more often.

Would the results be similar if the musical style would be changed from dodecaphony to a more tonal style, like renaissance counterpoint or choral music? Adapting Compunist to create music in a different style can be done, even though it will become a much more complex program.

A next step could also be to repeat this experiment, not using midi files, but having musicians play both the composed and generated music. It would be interesting to see whether the results are different. One has to take into account, however, that adding the human factor could potentially influence the results. With hard work and a lot of compassion, a musician can make a bad piece sound better. The opposite is true as well. A musician who doesn't care will sound boring and uninspiring. There would be the danger that the musicality or the compassion of the performer is measured, instead of the character of the piece itself.

10 Bibliography

Referenties

[1] Website Lilypond, http://lilypond.org/, last viewed on August 17, 2010.

[2] Website ThesisTools, http://www.thesistools.com/, last viewed on August 17, 2010.

[3] Website Finale, http://www.finalemusic.com/, last viewed on August 17, 2010.

[4] Website Sibelius, http://www.sibelius.com/, last viewed on August 17, 2010.

[5] Website about creating midi files http://www.kainhofer.com/ lilypond/ajax/user/lilypond/Creating-MIDI-files.html, last viewed on August 17, 2010.

[6] Automatic generation method of twelve tone row for musical composition used genetic algorithm, Maeda, Y. Kajihara, Y. Dept. of Human and Artificial Intell. Syst., Univ. of Fukui, Fukui, Japan, ISBN: 978-1-4244-3596-8

[7] Computer music composition systems questionaire, Peter Desain and Henkjan Honing, 1992.

[8] Compositietechnieken in de twintigste eeuw, Wim de Ruiter, De Toorts B.V., 1993.

11 Appendix

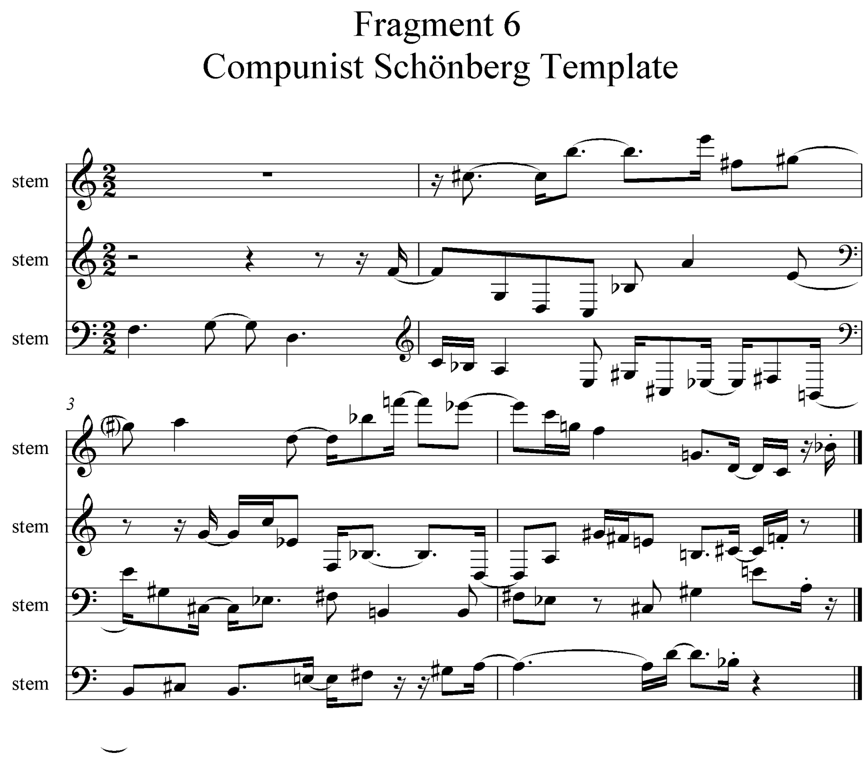
Fragment 2 (Schönberg Template)

Fragment 4 (Webern Template)

Fragment 5 (Stravinsky Template)

Fragment 5
Compunist Stravinsky Template

Fragment 6 (Schönberg Template)



Fragment 8 (Stravinsky Template)

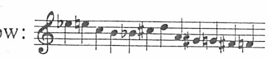
Fragment 9 (Webern Template)

Fragment 1: Kinderstück by Anton von Webern

2

PREFATORY NOTE

The *Kinderstück* was written by Webern in 1924, the first year in which he employed in his own music the twelve-tone method that Arnold Schoenberg had formulated the year before. Originally, Webern intended to compose a whole cycle of such pieces for young musicians, but he actually wrote only the present one. The little composition, long lost from view, came to light twenty years after Webern's death when, in October 1965, Hans Moldenhauer discovered it in manuscript, together with other unpublished works by the composer. According to Dr. Moldenhauer's printed catalogue of the Webern Archive*, the hunt for a sculpture representing Webern, "known only through a photograph, led to a dark attic in an old house near Vienna where, entirely accidentally, remnants of Webern's library and other belongings were uncovered along with the portrait bust. Among the relics were found many 'missing links' in Webern documentation, beginning with his earliest attempts at composition in 1899 and extending over the entire period of his creative work up to 1925."

The *Kinderstück* is based on the following elegantly constructed tone row: 

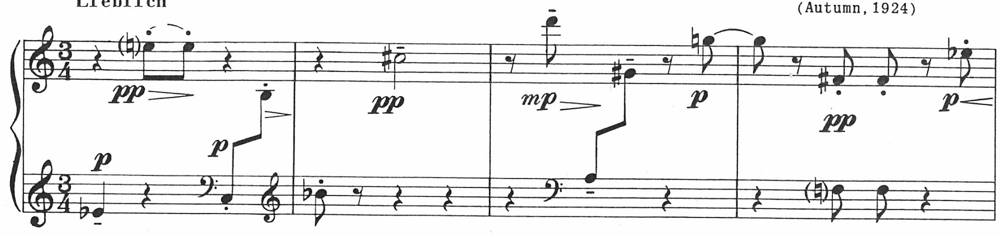
These twelve tones, always retained in their basic sequence (except where two or more tones are sounded simultaneously), can be traced, horizontally or vertically, throughout the piece. (In measure 4 from the end, the E-flat is written as a D-sharp.) In early works to which the twelve-tone principle was applied, such as this composition, all twelve notes of the chromatic scale had to be introduced before any of them could be sounded a second time, with one exception: a tone could be repeated immediately. This practice produces the "Morse code" effect peculiar to many early compositions in the idiom, and it is evident in this one.

The *Kinderstück* had its first public performance on July 22, 1966, at a concert in a Stravinsky Festival that was presented by the New York Philharmonic and that included some music by Stravinsky's more prominent contemporaries. On this occasion, the piece was played by Caren Glasser, aged nine. It is being published for the first time in the present edition.

Kinderstück

ANTON von WEBERN
(Autumn, 1924)

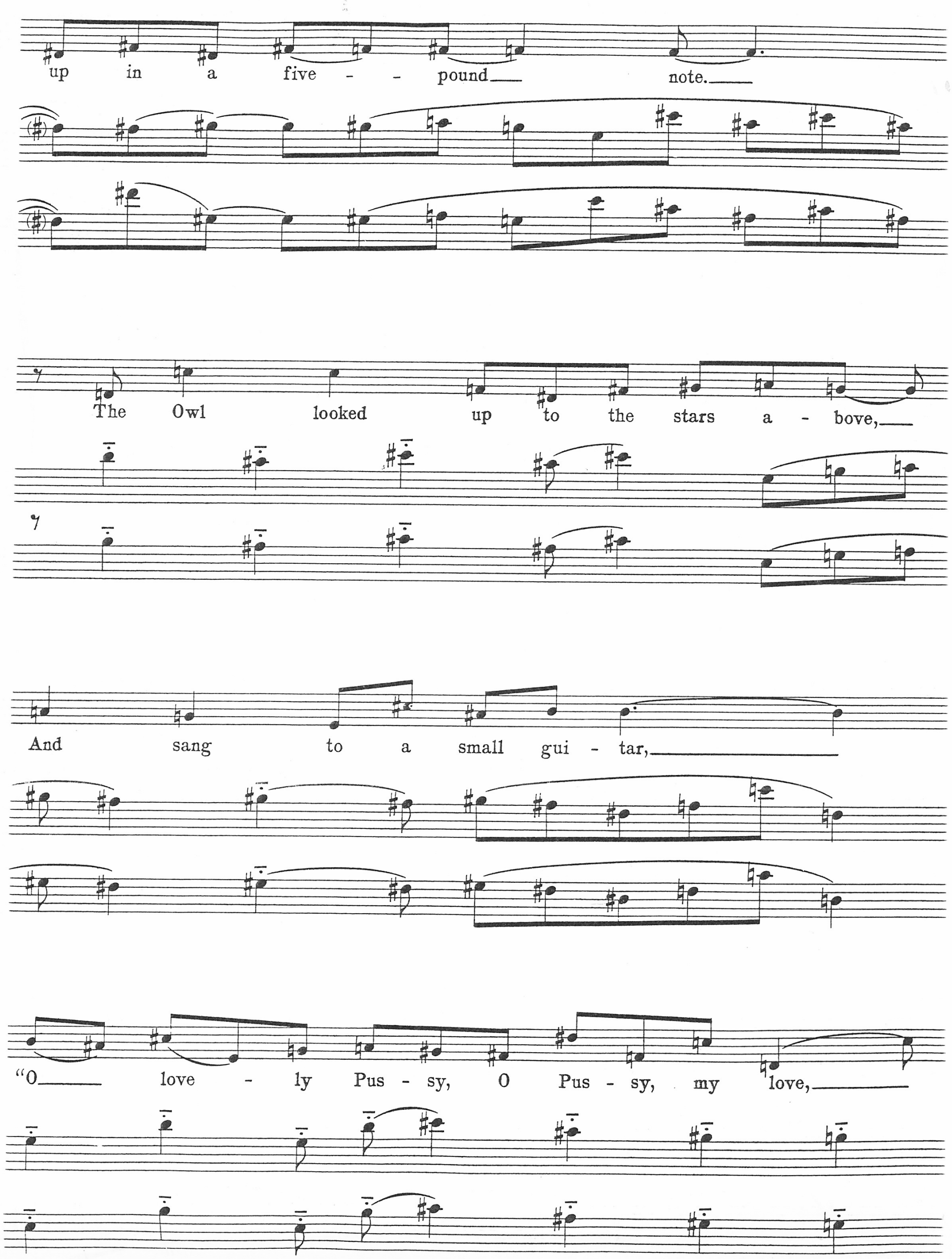
Lieblich



*In *Anton von Webern: Perspectives*, compiled by Hans Moldenhauer, edited by Demar Irvine, 1966, University of Washington Press.

B.& H. 19714

Fragment 3: The Owl and the Pussycat by Igor Stravinsky (part)


up in a five - - pound___ note.___
The Owl looked up to the stars a - bove,___
And sang to a small gui - tar,___
“O___ love - ly Pus - sy, O Pus - sy, my love,___


What a beau - ti - ful Pus - sy you are, You are,
You are! What a beau - ti - ful Pus - sy you are!"
Pus - sy said to the Owl, "You e - le - gant
fowl! How charm - ing - ly sweet you sing! Oh let us be

Fragment 7: Gavotte from opus 25 by Arnold Schönberg (part)


10
f
p
p
p dolce
p dolce
dolce
rit. . .
dim. . . . .
15
p
dolce
Ped. . . . . . *
f
tempo
5
4
p

Questionnaire

Enquête Preview

Enquête maken in 4 stappen:

Stap 1: *Enquete Instellingen*
Stap 2: *Enquete Overzicht*
Stap 3: *Enquete Preview*
Stap 4: *Enquete Publiceren*

Homepage: *Hoofdmenu*
Extra: *Extra diensten*

Page: 1

Art or artificial?

For my Bachelor project of the Leiden University, I'm researching whether people can hear the difference between computer generated and human composed twelve tone music.
Your cooperation is very much appreciated!

In this questionnaire you are asked to listen to some audio fragments.
For each audio fragment you are then asked to answer whether you think it was computer generated or human composed. You are also asked to say how sure you are about your answer, and for what reasen you have chosen the answer.

The fragments may sound unusual to you, if you never listen to contemporary classical music. However, it is very interesting whether a layman can hear the difference between the human and the computer fragments.
It's alright if you are not sure about your choice.

The questionnaire consists of 9 audio fragments of about 20 seconds each. Filling it in takes about 10 minutes.

Start

Page: 2

Art or artificial?

1.

What is your musical background? *

- Layman (knows very little about classical music)
- Listener (listens to classical music on a regular basis)
- Practitioner (plays an instrument)
- Professional (studies at a conservatory or is a graduate already)
- Expert (is a composer or a conductor, or listens to twelve tone music regularly)

Next, you can listen to the nine audio fragments.
Listen very carefully to each fragment.
After listening you are asked whether you think this fragment was randomly created by a computer, or composed by a human being.
It is very much appreciated if you also give the reason of your choice.
Each fragment takes about 20 seconds.

Fragment 1

3.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

4.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

5.

Waarop heeft u uw antwoord gebaseerd?

Next fragment

Page: 3

Fragment 2

7.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

8.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

9.

What did you base your answer on?

Next fragment

Page: 4

Fragment 3

11.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

12.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

13.

What did you base your answer on?

Next fragment

Fragment 4

15.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

16.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

17.

Waarop heeft u uw antwoord gebaseerd?

Next fragment

Page: 6

Fragment 5

19.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

20.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

21.

What did you base your answer on?

Next fragment

Page: 7

Fragment 6

23.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

24.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

25.

What did you base your answer on?

Next fragment

Page: 8

Fragment 7

27.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

28.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

29.

What did you base your answer on?

Next fragment

Page: 9

Fragment 8

31.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

32.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

33.

What did you base your answer on?

Next fragment

Page: 10

Fragment 9

35.

Do you think this fragment is randomly generated by a computer, or composed by a human being? *

- Created by a computer
- Composed by a human being

36.

How sure are you about your answer? *

- It's a guess
- Quite unsure
- Rather sure
- (Almost) sure

37.

What did you base your answer on?

Next page

Page: 11

Now please answer some questions about your background.

38.

Are you a man or a woman? *

- Man
- Woman

39.

How old are you? *

- 1-10
- 11-20
- 21-30
- 31-40
- 41-50
- 51-60
- 61-70
- 71-80
- >80

40.

Did you ever hear about twelve tone music before this questionnaire? *

- Yes
- No

41.

Do you live in a village or in a city? *

- Village (< 40000 inhabitants)
- City (> 40000 inhabitants)

42.

Do you have remarks about this questionnaire?

Submit!

Page: 12

Thank you for participating in my Bachelor research! The thesis with the results and conclusions will be available on this website from the beginning of September.