MusiCTM: On Music Composing by CTM

Zongyuan He Shuchen Li Chang Wang

August 8, 2021

Abstract

Creativity is a powerful advantage for human beings. We try to discover its interaction with consciousness by focusing on a specific creating process, music composing. Music composing was described as a three-phase process by Aranosian [1], that is, raw idea generation, formalization and development. We formulate these phases in the language of CTM (Conscious Turing Machine) and point out that consciousness and unconsciousness are both important in composing. In the first phase, flow state and associative thinking will help produce raw ideas unconsciously. In the second and the third phases, short-term memory and long-term memory interact, formalizing and learning the raw ideas, then determining what should follow, where consciousness involves. Finally, we hope that our work could inspire machine composing and extend to other forms of creation.

1 Introduction

Composing music has been thought of one of the most exclusive creative activities human beings can do. Surprisingly, machines now can compose music pieces that are even indistinguishable from human works in some sense (see [5] for a survey). Hence the problem on the difference of human and machine composers naturally arises.

In this project, we propose that consciousness is the key to this question. To illustrate this, we make an analogy with theorem proving mentioned in class. We think there are some similarities between theorem proving and artistic creation. They both require learning relevant knowledge before starting to create. And creative inspiration seems as if often coming from unconsciousness (like in the example of Poincaré). However, to complete the whole work, consciousness is also important. For example, when proving a theorem, one may come up with a high level idea unconsciously, and fill the dirty details, where consciousness involves. Also, in musical creation, instrumentation and mixing are more like engineering, which quite require consciousness. We study this process through the lens of *Conscious Turing Machine* (CTM, [4]).

2 Related Works

There is a line of works concerning music composing and consciousness. Employing the function of consciousness in music education has been studies extensively [12, 1]. In the latter paper, Aranosian argues that consciousness is a primary source for the ideas from which music is created. Sarath [15] also reveals the improvisation-driven growth of creativity and consciousness in the particularly improvisation-dominated form of music — jazz. On the other hand, Wiggins [18] presents a refinement of Baars' *Global Workspace Theory* [2] and uses the metaphor of the mind's chorus to account for the kind of creativity and inspiration *before* consciousness. He also extends the model [19] by unifying the account of non-conscious cognitive processing with conscious awareness. Beynon [6] also studies music composing from the perspective of computing and consciousness. The investigations into the evolution of music may also be of interest since they often address the effect of music on consciousness [13].

3 Background and Problem Formulation

Music is one of the universal cultural aspects of all human societies. However, the definition and creation of music may vary according to culture and social context. Some avant-garde works (like John Cage's *4'33"*), fringe styles (like hardcore punk), and random/casual creations (like jazz and humming) have been criticized as "not being music" throughout history. Generally, we define that music is the art of arranging sounds in time to produce an impromptu or permanent composition.

We distinguish an impromptu from a permanent composition here, based on the observation that different styles and types of music may emphasize or de-emphasize different music creation techniques. For example, classical music often requires several passes of careful organizing and fine polishing in a long period of music creation, while in jazz, improvising is heavily employed. It seems that in the former process there's more consciousness involved. As a starting point, we would like to extract a general model of music composing, provide simple explanation of this contrast, and further use the model to account for the role of consciousness in music composing.

Aranosian [1] describes musical composing as a three-phase process. The composer first comes up with a raw auditory idea, in the shape of captivating fragments. Secondly, the composer learns and expresses the idea in a concrete form, and places it in the memory. Lastly, the composer considers the development, i.e., what phrase should follow. The third phase will lead back to the first phase again. Thus, the whole process is like a chain of cause-effect or question-answer, in a cyclic manner. We point out that, in the first phase, the composer conceives the raw idea from unconsciousness, since many composers reported that they coming up their idea at odd, unpredictable times [3]. In the following sections, we will discuss how CTM can do these works.

4 The Process of Music Composing by CTM

If we recall the definition of music from section 3 — "music is the art of arranging sounds in time to produce an impromptu or permanent composition" — we can see that one must employ two mental channels when creating music; one channel arranges the elements of sounds, while the other channel is responsible for directly producing music — writing the musical notes down, playing an instrument, or conducting an orchestra. All of the assertions made about music above can also be applied to lingual activities, such as poetry and address making. Generally speaking, the employment of the two channels allows the creator to develop ideas and execute them at the same time, though specific functions of the channels vary among types of musical creation.

If we treat the two channels from the perspective of CTM dynamics, we can know that the chunks of ideas are created and refined by the first channel and then uploaded to the STM, then the STM finishes the work of the second channel — caching the idea flow and delivering it to external outputs. The two channels emerge alternately, but the caching mechanism of STM allows the continuous expression of music — a player don't usually pause when reading the next line on the staff.

In the following subsections, we will discuss what the first channel would be like in detail when processing the most challenging type of musical creation — composing. As we have noted above, the individual who makes a musical statement must initiate it, develop it with certain goals in mind, and conclude it, descriptions of the act of composing music mention a three-phase process [1].

4.1 Three-Phase Formulation

As we have pointed out, the first phase is with unconsciousness, so we would like to defer the discussion of this phase to subsection 4.2 and assume that a certain raw auditory fragment is placed into the STM when the second phase begins.

In the second phase, one thing the composer must do is to detect the pitches and durations of each note in the raw auditory idea, otherwise the composer cannot write down or play the piece he/she creates. Specifically, the composer will detect the relative pitch and relative duration rather then the absolute value, since the absolute values may be shifted (for the pitch) or scaled (for the duration) in the future arrangement. Besides, the composer will have a preliminary evaluation on the piece he/she generated. If the composer is satisfied with his/her idea, then he/she will move to the third phase. Otherwise, if the composer is not satisfied, but fortunately he/she knows how to fix it, then he/she will try to update this piece. If unfortunately the composer doesn't know how to fix it, then probably he/she will throw this idea away and go back to the first phase to seek some new ideas. However, we don't know the order of detection and evaluation, and we guess that it depends on the habit of the composer. In the language of CTM, this phase can be described as follows. We assume that the raw auditory fragment can be represented in one gist and there are detection LTM, evaluation LTM and fixing LTM, which will do the corresponding works. Without loss of generality, we assume that this composer detects before evaluates. First, the STM broadcasts the raw auditory fragment to all the LTM processors. On receiving the fragment, the detection LTM will generate a list of relative pitches and durations of the notes in the fragment (we also assume that this list can be represented by a single gist), with a relatively high weight so that the gist can reach the STM. Then the STM broadcasts these notes to the processors. This time, the evaluation LTM will decide whether this piece is satisfying. If it is, then generate a gist encoded with "satisfied", otherwise "not satisfied". Also it will have a relatively high weight so that it can reach the STM. The STM will broadcast this gist, and if it is "satisfied", this phase will finish. Otherwise the fixing LTM will generate a fix of the original piece (which it have stored in the memory in previous step) and send it to the STM. The evaluation process will run again. If still not satisfied, then the fixing process will run again. These two processes will run multiple times, until the evaluation LTM is satisfied or the running time has reached a certain threshold. In the latter case, the piece will be thrown away and the CTM will go back to the first phase. Figure 1 shows an example (for simplicity, we omit the note duration and some irrelevant broadcast, and the piece is very short).

After the composer has settled an idea down, he must determine how to complete his work from this small music piece. This is when the third phase starts, and in this phase consciousness is heavily involved. This part with consciousness is indispensable because generally the composers cannot rely solely on ideas, as Stuckey [16] points out that one cannot make music with ideas, or poetic dreams or wishful thinking and it is technique that paints paintings, writes poems, builds buildings. Typically the settled idea is a motif in classical music, and there are many existing ways to extend it. The composer may want to choose some forms of music, like rondo, to pin down the general framework of the whole piece. Furthermore, standard methods including progression, variations, counterpoint and modulation could be adopted to develop the music. In the language of CTM, this process can be described as follows. The CTM puts a single strong signal, consist of the motif and the question of how to proceed with the motif, and then the signal gets broadcast. The LTMs that are able to provide possible solutions then start working by recognising this broadcast. They may estimate the likelihood of success by using the technique the LTM correponds to, like counterpart, determine the weights and report. We note that there will be few LTM processors that are able to clearly

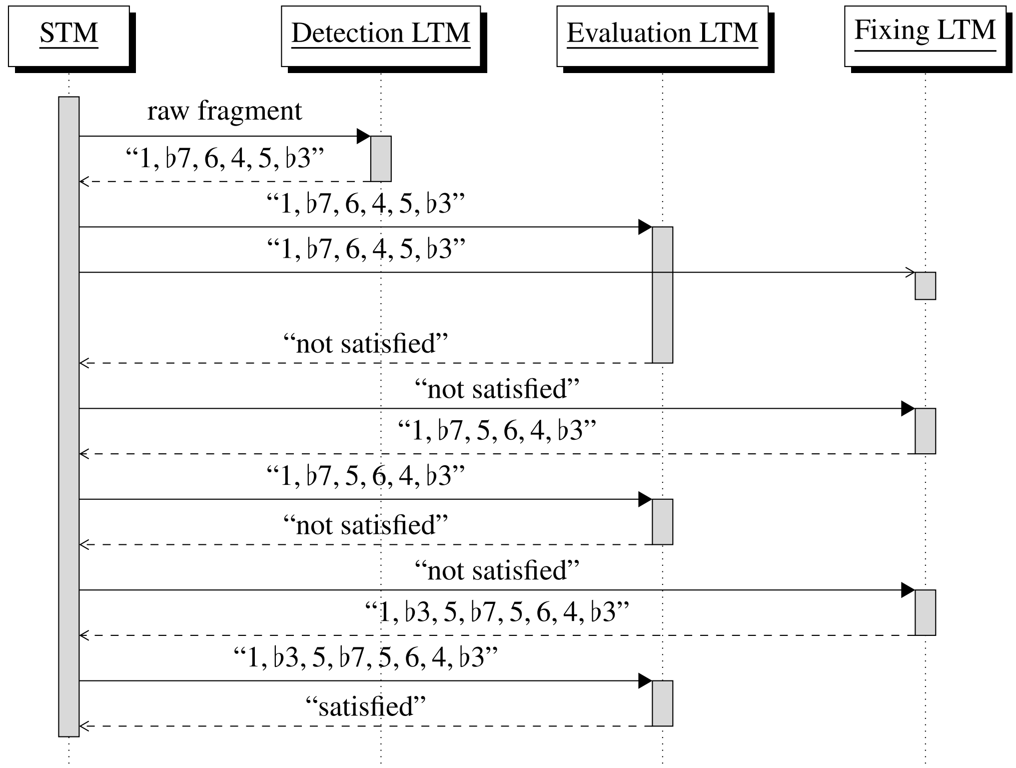

Figure 1: An example for the second phase

recognise this instruction, and since they are specialized to do such narrowed processing work, there's hardly any competition through the up tree. Therefore, the result should be clear and intense. This is why we can proceed very fast once we get some ideas for composing. Having done the third phase, the composer makes a step by getting a hint from one LTM, after which he returns to the first phase and again seeks for (more targeted) inspirations. In summary, the appropriate composing templates or routine methods are invoked through consciousness, which will provide useful hints for phase one of the next iteration and clearly boost the whole process of composing. Figure 2 shows an example.

4.2 Raw Idea Generating

As Margery Vaughan [17] pointed out, associative thinking is a key to creativity. It may be an important source of raw ideas. For example, composers may come up with a raw idea when listening to other musical pieces, reading an essay, enjoying the scenery, or experiencing some emotions. This is a case of *Synesthesia*. Synesthesia is a perceptual phenomenon in which stimulation of one sense leads to another sense. Awareness and type of synesthesia varies from person to person. In a most common type of synesthesia, people may perceive letters with certain colors for each letter, which is called grapheme–color synesthesia. The mechanism of synesthesia is still conjectured, but there are two major theories, *cross-activation* and *disinhibited feedback*, have been proposed. They both start from the observation of the dedicated regions of the brain. Cross-activation theory suggests that the increased cross-talk between different dedicated regions may account for some types of synesthesia. For example, the grapheme–color synesthesia may come from the cross-activation of the grapheme-recognition area and color area called V4 [14]. An alternative

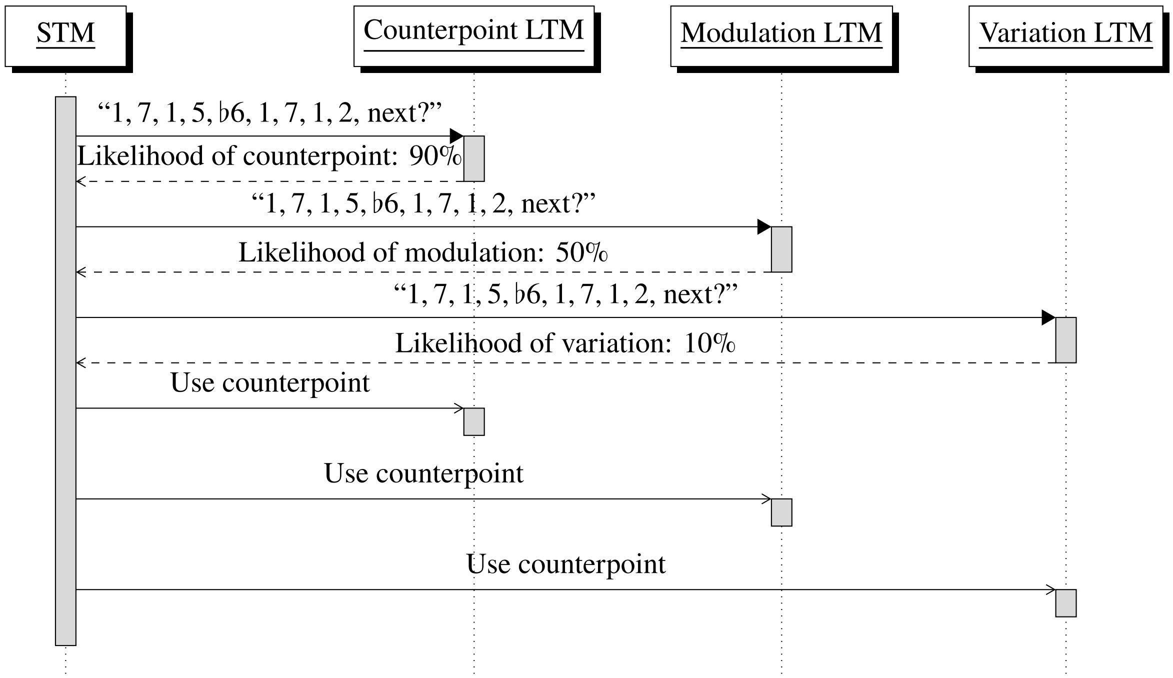

Figure 2: An example for the third phase

theory suggests that disinhibited feedback may be the cause. That is, if a feedback is not inhibited (as usual), then the feedback signals from late stages of multi-sensory processing might influence earlier stages [11]. Both theories involve interactions among the dedicated region of the brain. By analogy, we can interpret the LTM processors as the dedicated regions of brain, and the “links” between LTM processors as the interactions between different regions of brain. For example, when enjoying the scenery, the visual information processing LTM will send relevant information thought the link, to a raw idea generating LTM. A well-trained composer will have strong links between these LTMs. After receiving enough information, the raw idea generating LTM will generate a gist of a raw auditory fragment.

However, this is not sufficient: the weight of the gist may be too small to reach the STM. So only in certain mind states can people start to compose. As Wiggins quoted Mozart’s introspective account of his creative process [18], Mozart need to be content and undistracted to become conscious of musical ideas emerging from somewhere. The musical idea is more likely to fail in the up-tree competition in the presence of distractions, such as hunger, communication with others, or finding a route. Then there’s no distractions, the chunk of musical idea will have a relatively higher weight, ans thus is more likely to reach the STM. Only in such case can the composer be conscious of the raw auditory fragment, and move onto the second phase as described in subsection 4.1.

> When I am, as it were, completely myself, entirely alone, and of good cheer – say traveling in a carriage, or walking after a good meal, or during the night when I cannot sleep; it is on such occasions that my ideas *flow* best and most abundantly. Whence and how they come, I know not; nor can I force them. Those ideas that please me I retain in memory, and am accustomed, as I have been told, to hum them to myself.
> All this fires my soul, and provided I am not disturbed, my subject enlarges itself, becomes methodised and defined, and the whole, though it be long, stands almost completed and

> finished in my mind, so that I can survey it, like a fine picture or a beautiful statue, at a glance. Nor do I hear in my imagination the parts successively, but I hear them, as it were, all at once. What a delight this is I cannot tell! All this inventing, this producing takes place in a pleasing lively dream. What has been thus produced I do not easily forget, and this is perhaps the best gift I have my Divine Maker to thank for.
> (W. A. Mozart, quoted by Holmes [9])

As we noted, under certain circumstances, the creator would be completely absorbed in what he or she does, i.e., the creator is fully immersed in a feeling of full involvement, energized focus, and enjoyment in the process of the activity, to reach *flow state* (a mental state named by Mihaly Csikszentmihalyi [8]). Flow state is believed and proved to promote musical creativity [10]. Mozart's telling of his own experience demonstrates this state and its emergence from a composer's perspective. Csikszentmihalyi [7] outlined specific features of flow which may ensure that activities can be both enjoyable and rewarding for participants. For example, providing 'clear goals every step of the way' and instant feedback on the performance during the activity, ensuring that 'there is no worry of failure', and containing a balance between challenge and skill, these features often provide individuals with the exhilarating feeling – flow.

We call it a flow state of CTM when LTMs with similar functions communicates repeatedly with the STM, they "resonate" together and the weights of these LTMs gets dominant, meanwhile the weights of LTMs with conflicting or opposite functions are diminished. When a creator is in the state of flow, the LTMs responding actively to physical discomforts, changes of surrounding environment, and the possibility of failure are relatively quiet to the STM. If the flow gets interrupted, the weights of involved LTMs in the flow state would decay over time, so a balance should be kept between the goal or perceived challenge and the individual's skill. Immediate feedback is also necessary for the maintenance of the state so that the melody could get broadcasted to LTMs. A composer usually needs a piano to make the notes in mind audible to finish the composition. In his later deaf years, Beethoven made use of a wooden stick transmitting the vibration of sound to his skull to compose.

5 Discussions

Having completed the discussion of the model, we aim at discussing some examples to show the strength of our theory. The first example that we would like to address is the difference between improvising and composing. When a musician is improvising, the development of music is crucial for the success. He must determine what to play next almost instantly. Hence he cannot place the melody he just played into STM, nor does he have any time to make repeated assessments of existing ideas. According to the three-phase formulation, it's phase two's job to evaluate and settle down an idea, and in phase three the following pieces of music are generated and signalled in a more deterministic manner. Therefore, in improvising phase two is de-emphasized and even skipped. Furthermore, by exploiting the links between the LTMs, the following pieces of music can be generated quickly, without using consciousness. This is why one must be specially trained and accumulate lots of resources before he can improvise well.

Secondly, we would like to answer the question mentioned at the beginning of the paper, concerning the differences between human and machine composers. As is promised, we explain this in terms of consciousness and unconsciousness.

comprehensive and holistic thinking In CTM, there are sufficiently large number of LTMs, but for machines, they have only limited number of memory and processors. They can memorize composed excerpts through techniques like LSTM, at best. Therefore, when machines are composing music, they can only take a small piece of music into consideration and append without a holistic view. The algorithm concentrates on *local* melodies and is, roughly speaking, sequential. This fundamental difference shows that although machines may compose short pieces that sound natural, they cannot compose long and comprehensive symphonies with high quality.

complex unconscious links As is mentioned above, when composers are seeking for raw ideas, generally he is not intentionally targeting at composing a piece of music. A gist generated by some totally irrelevant LTM may get directly transmitted to other LTMs without passing through the STM. The idea can also propagate to be transmitted further, making influences on a large mesh or network of LTMs. Thus any interesting experience can be a source of music composing. An extreme phenomenon is called *synesthesia*, a perceptual phenomenon in which stimulation of one sensory or cognitive pathway leads to involuntary experiences in a second sensory or cognitive pathway. In the field of music, for example, Alexander Scriabin composed colored music that was deliberately contrived and based on the circle of fifths, whereas Olivier Messiaen invented a new method of composition (the modes of limited transposition) specifically to render his bi-directional sound-color synesthesia, such as the red rocks of Bryce Canyon are depicted in his symphony *Des canyons aux étoiles*. This example demonstrates that the enormous number of links among the LTMs enable humans to do more creative work and make avant-garde music. Humans definitely do not limit themselves to datasets as machines do.

conscious evaluation and dynamical updating Following the discussion of complex unconscious links, we must also talk about how these links are formed, which involves consciousness. Recall that composers make evaluations in phase two and fix the ideas that he's not satisfied with. The LTMs observe the content inside the STM and make assessments of its reported gist (idea). It then tries to update his weights and to make some links to other more successful LTMs. This dynamical updating process can be implemented in machines, but machines can only evaluate with some fixed and fine-tuned rules and is merely able to self-evaluate. Furthermore, human composers only need a few shots of instances to learn to reflect and update. The learning process is also online rather than offline.

To summarize, we have presented a theory in the language of CTM to account for the process of music composing. We also make some interesting observations from it.

References

[1] Christopher M Aranosian. "Musical Creativity: The Stream of consciousness in composition, improvisation, and education". In: *Imagination, Cognition and Personality* 1.1 (1981), pp. 67–88.

[2] Bernard J Baars. "In the theatre of consciousness. Global workspace theory, a rigorous scientific theory of consciousness". In: *Journal of consciousness Studies* 4.4 (1997), pp. 292–309.

[3] Stan Bennett. "The Process of Musical Creation: Interviews with Eight Composers". In: *Journal of Research in Music Education* 24.1 (1976), pp. 3–13.

[4] Manuel Blum and Lenore Blum. "A Theoretical Computer Science Perspective on Consciousness". In: *Journal of Artificial Intelligence and Consciousness* 8.1 (2021), pp. 1–42.

[5] Jean-Pierre Briot, Gaëtan Hadjeres, and François-David Pachet. "Deep Learning Techniques for Music Generation – A Survey". In: *arXiv preprint arXiv:1709.01620* (2017).

[6] David Clarke and Eric Clarke. *Music and consciousness: Philosophical, psychological, and cultural perspectives*. Oxford University Press, 2011.

[7] M. Csikszentmihalyi. *Flow: The Psychology of Happiness*. Random House, 2013. ISBN: 9781448177707.

[8] M. Csikszentmihalyi and I.S. Csikszentmihalyi. *Beyond Boredom and Anxiety: Experiencing Flow in Work and Play*. Jossey-Bass Publishers, 1975, pp. 10–. ISBN: 9780787951405.

[9] Edward Holmes. "Part the Fourth: His Manhood". In: *The Life of Mozart: Including his Correspondence*. Cambridge Library Collection - Music. Cambridge University Press, 2009, pp. 317–318.

[10] Raymond MacDonald, Charles Byrne, and Lana Carlton. "Creativity and flow in musical composition: an empirical investigation". In: *Psychology of Music* 34.3 (2006), pp. 292–306.

[11] "Mechanisms of synesthesia: cognitive and physiological constraints". In: *Trends in Cognitive Sciences* 5.1 (2001), pp. 36–41. ISSN: 1364-6613.

[12] Anthony J Palmer. "Consciousness Studies and a Philosophy of Music Education". In: *Philosophy of Music Education Review* (2000), pp. 99–110.

[13] Leonid Perlovsky. "Musical emotions: Functions, origins, evolution". In: *Physics of life reviews* 7.1 (2010), pp. 2–27.

[14] V. Ramachandran and E. Hubbard. "Synaesthesia – A Window Into Perception, Thought and Language". In: *Journal of Consciousness Studies* 8.12 (2001), pp. 3–34.

[15] Edward W Sarath. *Improvisation, Creativity, and Consciousness: Jazz as Integral Template for Music, Education, and Society*. SUNY Press, 2013.

[16] Steven Stuckey. "Creating music of geometry and longing". In: *Cornell Univ Arts Sci Newsl* 18.1 (1996).

[17] Margery M. Vaughan. "Cultivating Creative Behavior: Energy Levels and the Process of Creativity". In: *Music Educators Journal* 59.8 (1973), pp. 35–37.

[18] Geraint A Wiggins. "The Mind's Chorus: Creativity Before Consciousness". In: *Cognitive Computation* 4.3 (2012), pp. 306–319.

[19] Geraint A. Wiggins. "Creativity, information, and consciousness: The information dynamics of thinking". In: *Physics of Life Reviews* 34-35 (2020), pp. 1–39.