Considering the Interaction of Humans and Music from a Human-Agent Interaction Approach

Benjamin Luke Evans
Graduate School of Information Science and Technology, Hokkaido University
benjamin@complex.ist.hokudai.ac.jp

Nagisa Munekata
Graduate School of Information Science and Technology, Hokkaido University
munekata@complex.ist.hokudai.ac.jp

Tetsuo Ono
Graduate School of Information Science and Technology, Hokkaido University
tono@ist.hokudai.ac.jp

ABSTRACT
The study of automatic music generation has been an interest of scientists from many fields, leading to a diverse assortment of studies and systems which incorporate different methods and technologies. However, it is noticeable that this diversity has led to many systems today focusing on the quality of individual performances and that the general study of automatic music generation lacks the framework for systematic discussions. In our study, we have taken the scope of automatic music generation research that includes human interaction in the system design. We have taken models of human-agent interaction and applied them to human-music interaction and have considered a new generalized framework that would lead to constructive discussions in this context. In particular, we point out a lack in the definition of the "environment" automatic music composition systems are aiming to affect and start to consider some guidelines to be used in future research.

Author Keywords
Music; human computer interaction; agents.

INTRODUCTION
The effects of music on listeners have interested musicians and philosophers for centuries and have greatly influenced the formation of popular music today. Science has taken notice particularly of the emotions conveyed through music and empirical studies have been conducted for more than a hundred years [1]. Recently, with the rise of technology represented in the development of the internet, research has taken notice of the social relations between humans centered around music, which have shared in the development of new musical services, such as *Songle* [2] and *Songrium* [3].

A noticeable trait of music-information research today is that many researches focus on the quality and design of the individual system and lack a systematic discussion of the general field. Many researches are self-conclusive and are weak in generalized contributions to other researches. A typical example of this can be seen in the studies of automatic-contribution systems. Many systems have been developed using different algorithms and computational techniques (see [4] for details), but hardly any systematic information has been attained from these that contribute to the foundation of development of future automatic composition systems.

To address this problem, we have borrowed the analogy of "intelligent agents", as described by Russell and Norvig [5], and have defined automatic music composition systems as "agents". We have taken research models of human-agent interaction and have applied those to human-music interaction. By using a predefined model used in another field of research, we hope to introduce an objective framework for discussing automatic music composition systems in general and a means to lay constructive contributions for future system developments.

MUSIC COMPOSITION SYSTEMS AS AGENTS
Russell and Norvig [5] describe agents as "anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators." From this description, we derive three points for which something can be identified as an agent.

1) It must be conceivable as an independent entity.
2) It must perceive its environment through sensors.
3) It must affect its environment through actuators.

Below, we will discuss how automatic music composition systems fulfill the given criteria and therefore can be described as agents.

Conceivable as an independent entity
Music composition systems are easily conceivable as independent entities. This is evident in the way researchers have developed systems that incorporate specific methodologies and can be identified as differing from each other. (See [4] for summery of different automatic composition systems.) Furthermore, most of these systems are given unique names promoting the fact that they are "independent" and indeed self-contained "entities".

Perceives environment through sensors
[5] gives an example of software agents receiving keystrokes, file contents and network packets as sensory input. In general, as most automatic composition systems are software, these same sensory inputs can be applied here as well. File contents and network packets may include, but are not limited to, the following: Biophysical data of users, audio information of physical environment, tastes and trends deducted from previous user experiences, possibilities/limitations given by the choice of instrument/device used in audio feedback, etc. Theoretically, automatic composition systems that are

initialized by a random seed and conclude with absolutely no user intervention could exist. However, a vast majority of systems do include user-input in one way or another and therefore can be said to "perceive its environment".

Affects environment through actuators

In the case of automatic composition systems, the actuator through which it affects its environment would be the produced music. More specifically, the continuation (and resulting discontinuation) of sounds is what it affects the environment with.

As described above, automatic music composition systems do indeed fill the criteria of an agent derived from [5]. It is quite conceivable to consider an automatic composition system as an agent. However, one question does remain; What is this agent's environment? We discuss this in the following section.

THE ENVIRONMENT OF COMPOSITION SYSTEMS

[5]'s description of the agent's environment is less defined than that of the agent itself. Ohuchi et. al [6] define the environment as all that not pertaining to the entity of the agent. Though theoretically valuable, this definition is too broad in the context of a real situation and lacks in practical application. Ohuchi et. al go on to say that a more limited environment must be defined for each agent application.

When considering the definition of the environment for automatic music composition systems, different scenarios appear depending on the direction of research. Some systems are developed to generate new sounds, or new combination of sounds, focusing on the aesthetic side of music. The environment in this case could be defined as the performance of the completed piece, the artistic appreciation of each user, or the listening audience as a whole. As can be seen, the definition of an environment, even in the context of a single automatic composition systems, can be very diverse.

[5] explains (task-) environments as being "the 'problems' to which rational agents are the 'solution.'" [5] also makes clear that the performance of an agent is measured by the evaluation of environment states, not agent states. Therefore, the definition of the environment (the problems a system is a solution to and the means by which its performance is evaluated) is crucial in both designing and evaluating agents. Yet, when closely reviewing studies pertaining to automatic composition systems, we notice most lack in the definition of their environments.

Automatic composition systems are often presented as an example of new algorithmic application or as an artistic product to be used for its aesthetic properties. Research is often focused on the details of the system (means of musical composition) or its actuator's actions (produced music) and lacks in considering the whole space to which the system will be applied. We consider this lack of definition regarding system environment contributing to the lack of systematic discussion in this field. A lack of definition of environment leaves each research no solution but to evaluate its success against itself, discouraging any formation of objective and systematic evaluation across systems. This self-evaluation is precisely what [5] warns about when describing performance measurement.

CONCLUSION

In this paper, we have successfully defined automatic music composition systems as agents, borrowing a typical model from the field of human-agent interaction. While the definition of agents is relatively clear in this model, the definition is made using a less-defined article, the environment. Though the importance of the definition of the environment for each application is recognized in the field of human-agent interaction, it is rarely practiced in studies of automatic composition systems. We have pointed out how this lack of definition of environment has led to each study's self-evaluation and the discouragement of formation in systematic arguments that contribute to other research.

In future, we plan to continue looking at music composition systems as agents and assess current studies of automatic music composition using models given in human-agent interaction research. We look to form new models of research specific to studies of automatic composition that will lead to the formation of systematic and objective share of information across multiple studies in this field.

REFERENCES

1. MacDorman, K. F. Automatic Emotion Prediction of Song Excerpts: Index Construction, Algorithm Design, and Empirical Comparison. *Journal of New Music Research*, *36*, 4 (2007), 281-299.
2. Masataka Goto, Kazuyoshi Yoshii, Hiromasa Fujihara, Matthias Mauch, and Tomoyasu Nakano. Songle: A Web Service for Active Music Listening Improved by User Contributions. In *Proc. ISMIR 2011* (2011), 311-316.
3. Masahiro Hamasaki and Masataka Goto. Songrium: A Music Browsing Assistance Service Based on Visualization of Massive Open Collaboration Within Music Content Creation Community. In *Proc. ACM WikiSym + OpenSym 2013* (2013), 1-10.
4. Benjamin Evans, Satoru Fukayama, Masataka Goto, Nagisa Munekata, and Tetsuo Ono. AutoChorusCreator: Four-Part Chorus Generator with Musical Feature Control, Using Search Spaces Constructed from Rules of Music Theory. In Proc. ICMC + SMC 2014 (2014), 1016-102.
5. Russell, S. and Norvig, P. *Artificial Intelligence: A modern Approach.* Pearson Education, Inc., Upper Saddle River, NJ, USA, 2010.
6. Ohuchi, A., Yamamoto, M. and Kawamura, H. *Theory and Application of Multi-agent Systems: Computing Paradigm from Complex Systems Engineering.* Corona Publishing, Tokyo, Japan, 2011