The Education of the AI Composer: Automating Musical Creativity

Benjamin D. Smith
National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801
Email: bdsmith3@illinois.edu

Guy E. Garnett
Illinois Informatics Institute
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

***Abstract*—Models of computational creativity promise to provide insight into the nature of human creative work and innovation. Intrinsically motivated, automated, creative agents present a potential avenue for the exploration of creativity in the arts, and in music in particular. A novel reinforcement learning agent designed to improvise music in a self-motivated fashion is described, formulated to prove the capabilities of an artificially creative musical system. The prototype employs unsupervised adaptive resonance theory algorithms to model theories of human perception, cognition, and creativity. While the generated results are constrained for initial evaluation further extensions suggest the potential to create meaningful, aesthetically valuable compositions.**

I. Introduction

The formulation of a creative, automatic musical composer, improviser, or companion is a current topic [1] with the potential to provide significant insight into *creativity, inspiration,* and *invention*. Machine learning applications, typically formulated around theories of biological brain function, seek to understand observed phenomena through pattern abstraction and reduction. Ideally these artificially intelligent systems produce useful conceptual models of the generative processes underlying the observed data such that the system can respond to new stimuli in a meaningful way. Yet, machine learning algorithms struggle to assess the external significance of new material or concepts on their own, and require careful formulation (by the designers) to produce significant results. Thus, these systems have no inherent sense of how original or valuable their creations are, especially in relation to the external, human world.

If an automated agent can be given a self-applicable, generic metric for the *creativity* of its actions it could evaluate its own ideas or discoveries in a useful fashion. Such a system could then produce results with a greater chance of being meaningful to an outside observer, rather than wandering randomly through its prescribed data or concept space. At the very least these automatically originated ideas would be more efficient and have internal value as they serve to map out the agents conceptual domain.

Reinforcement learning, a specialized machine learning model that enables an agent to explore, discover, and explain its own environment, presents possibilities in this direction. The typical reinforcement learner (RL) attempts to optimize its behavior towards predefined, externally determined conditions (such as successfully navigating a maze). However a RL can be adapted to model computational creativity [2] by providing the agent with an internal measure of *novelty*, allowing an artificial agent to actively create and/or discover patterns which have personal (for the agent) value [3], [4]. This intrinsically-motivated RL is able to evaluate its discoveries based on a sense of *novelty* or *surprise*, encouraging itself to identify new patterns and new algorithms that enable a better mapping (i.e. understanding) of its environment. Ideally the agent is able to further identify culturally or historically meaningful concepts, given sufficient background and context.

After a brief discussion of computational creativity, we present a novel RL design and prototype system created to improvise music alone or in an ensemble. Employing unsupervised machine learning algorithms, this RL intentionally models human musical creativity and inspiration at a fundamental level and is able to operate in live, performance settings.

II. Computational Creativity

Modeling creative processes in a computational system is inherently problematic. Notions of *creativity* seem to be intrinsically tied to human agency and determining the success of an automated system in an objective fashion may be futile. Boden [3] presents a useful dichotomy in framing the subjective nature of creativity: that of historically valuable innovation (H-creativity) versus personally or psychologically meaningful discovery (P-creativity). H-creative ideas have value to a larger community and may be truly original formulations, occurring for the first time in recorded history (such as the formidable works of the western master composers, Bach, Mozart, etc.). On the other hand, P-creative innovations are only required to have personal significance, i.e. they are novel for the individual only (which would include any musical creation which its author finds compelling). Certainly P-creative ideas may also be H-creative, but the later requires acknowledgement from a wider community, potentially all of humanity. Additionally there must be a continuum of H-creative value, from Nobel prize winning discoveries to a musical composition that is considered innovative within a small community of practitioners.

Creativity is a process of formulating new concepts that either combine or extend existing ideas. Semioticians argue

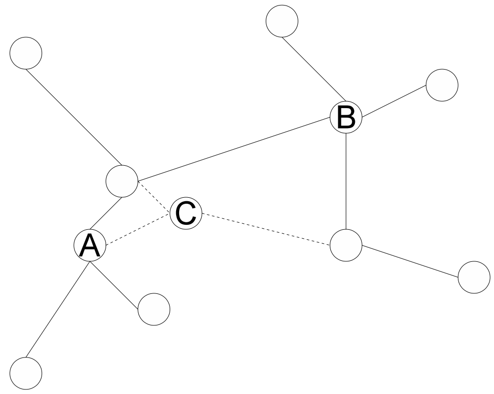

Fig. 1. Associative map.

that all human thought is built through associative maps, where the knowledge contained in any one idea is defined by its relative position (association) to every other idea. Figure 1 depicts such a map, in which ideas A, B, and C are related by their connections in the idea space. Combinatorial creativity is then understood as the creation of a new node in the associative map that lies between two known nodes, combining some elements of each to derive a third. Consider if node C in figure 1 were initially absent leaving only a single conceptual path between nodes A and B. Next, C is created, providing a new path relating A and B in an original way, for this map. The creation of C, if limited to the thinking of an individual, would be deemed P-creative as it is novel and changes the understanding of the individual in question. However if the map represents the collective thinking of a larger community it could be construed as H-creative, modifying the conceptualization of a field of thought.

A more transformative form of creativity is the discovery of new concepts that exist outside the bounds of the known map and may result in a reformulation of some or all of the map (called *exploratory creativity* [3]). These types of discoveries typically have a significant impact for the individual or community within which they are understood, such as the movement to serial tonality and the adoption of stochastic processes in music in the last century. These shifts in thinking transform the previously accepted idea spaces in radical ways, just as the recognition of a third dimension (extending up from the page) would suddenly transform figure 1 into a dramatically different associative map.

When evaluating acts of creativity, or computationally creative systems, one or all of the following four elements must be satisfied [from [4], originally describing human creativity in problem solving]:

1) The product of the thinking (system) has novelty and value (either P-creative or H-creative).
2) The thinking is unconventional, in the sense that it requires modification or rejection of previously-accepted ideas.
3) The thinking requires high motivation and persistence.
4) The problem-area is vague, so that part of the thinking is to formulate the problem itself.

The nature of the novelty of the product (element 1) may be deemed H-creative through some form of Turing test with human subjects, but P-creativity may be argued through proving element 2 (i.e. if the thinking requires the reformulation of the agents concept map through the modification or rejection of previously-accepted ideas, then it is deemed P-creative). The value assigned to the thinking may be in relation to another process (such as survival) or an external mandate (such as passing a test). In a computational sense element 2 can be understood as change in an associative map, adding or removing entries in a database that link to all other entries, thus modifying the meaning of all known data. Additionally, self-modifying systems, so long as the modification is novel, non-formulaic and valuable, may be deemed creative according to element 2.

High-motivation and persistence indicates the dedication of the agent to the task of discovery, as any other approach is likely to undervalue the innovation (due to ignorance or apathy). This does not deny spontaneous inspiration, as this is typically a result of sub-conscious work on a problem that is highly valued by the individual (such as Archimedes' famous "Eureka!" discovery). Element 4 acknowledges that creativity often appears in poorly defined areas of thought and research, where simply discovering the challenges and mapping the problem domain are significant contributions.

As long as working definitions of creativity (such as that proposed by [4], above) do not arbitrarily require human agency it is possible to consider automated creativity. Yet, is the presence of creativity a viable aesthetic measure? Conclusively answering such a challenge is likely infeasible, due to the subjective nature of aesthetics and the individual valuation of creative products. Yet conversely, by considering artistic products that are very low in creative content (i.e. they are predictable, conventional, created in a cursory fashion, and exist in a well defined area) it seems clear that a lack of creativity is an aesthetic liability.

III. Reinforcement Learning

Human creative agency is observable from birth and infant cognitive development serves as a clear model for general computationally creative, artificial agents. The intrinsically motivated RL model [2] is formulated as an extension of human learning processes, extending the conventional stochastic RL with internal metrics of agent creativity. While humans readily learn from sensory stimuli and their environment, avoiding heat, injury, hunger, and thirst, they take more than a passive role in this process. Babies actively conduct experiments of the nature "what sensory feedback do I get if I move my eyes or my fingers …just like that?" [2] In this way the individual is always seeking new effects that exhibit "some yet unexplained but *easily learnable* regularity." Stimuli observed previously is quickly deemed boring, while entirely new input is regarded as incomprehensible noise. Through this gradual mapping of behaviors and patterns the learner gradually acquires more and more complex behaviors, eventually leading to the extreme abstractions (in humans) of academic thought, scientific innovation, and aesthetic inspiration.

A simple algorithmic mechanism is proposed by [2] to explain this learning phenomena, which uses RL to maximize the "internal joy" of the discovery of "novel patterns." Patterns can be understood as regularities in a dataset that can be abstracted in some fashion and effectively reduced, in complexity or size, as a result (i.e. data compression). Thus as new inputs are observed and recorded the discovery of new patterns allows the observed history to be recorded in smaller and smaller memory spaces. When an agent discovers a regularity or a new model that allows phenomena to be compressed, the pattern is deemed temporarily *interesting*. This process can also be described as the refinement of an associative map, where each new pattern is a new node in the concept space.

The crucial elements of the intrinsically motivated RL model are:

1) An adaptive world model, essentially a predictor or compressor of the continually growing history of actions/events/sensory-inputs, reflecting what is currently known about how the world works,
2) A learning algorithm that continually improves the model (detecting novel, initially surprising spatiotemporal patterns that subsequently become known patterns),
3) Intrinsic rewards measuring the model's improvements (first derivative of the learning progress) due to the learning algorithm (thus measuring the degree subjective surprise or fun),
4) A separate reward optimizer or reinforcement learner, which translates those rewards into action sequences or behaviors expected to optimize future reward.

Building a musical RL requires 1) a model of music, a predictor/compressor containing all the music heard by the agent and containing everything that is "known" about music; 2) an algorithm that learns how music works (improves the model in 1); 3) a reward measure of the model's improvements in (2); and 4) an agent that creates more music anticipating maximal future reward. In other words, (1) is an application specific analysis of all the music presented to the agent and (2) is the set of working theories and concepts that explain these analyses.

To implement 1) and 2) we employ spatial feature encoding [5] and Adaptive Resonance Theory (ART) [6], an unsupervised machine learning model mimicking elements of human cognition and perception [7]. Element 3) is understood as a measure of the relative entropy between the ARTs prior and posterior states, diverging from conventional RL implementations which typically employ a random exploration model. Our curious exploration model [2] ensures that the music is always internally interesting and valuable (satisfying P-creative requirements). Finally, 4) is implemented as a comprehensive predictor that anticipates the intrinsic reward measure for every potential stimuli and acts to maximize this reward.

IV. Discussion

The ART employed in our RL is a generic fuzzy ART implementation [8], providing the capability of identifying significant patterns in a series of feature vectors and adaptively encoding them into a neural network. In order to model the Wundt curve of the agent (i.e. the spectrum between boredom and confusion) we take the amount of change in the ART network (the residual, r in equation 1) as a measure of *novelty*. Thus new inputs that produce more change in the system (both in terms of changing the weights of network nodes as well as involving new nodes) move towards the chaotic end of the continuum. Conversely, inputs that produce no change are deemed boring. Thus the agent's intrinsic reward (i) is a function of the change in the size (e) of the network (*relative entropy*, or how many new bits of data are required to store the network) modified by the amount of change seen in the node's weights. The Wundt curve is modeled by inverting the relative entropy such that a balance is found between modifying the weights of existing nodes and creating new categories (i.e. new nodes), where a middle ground is deemed to indicate "easily learnable" [2].

$$i = \frac{re_t}{e_{t-1}} \tag{1}$$

Additionally, the intrinsic value of a given input is modified by the resonance (a measure of proximity between the input vector and each encoded category, $\mathbf{R}$ in equation 2) from each node in the network. In this way new inputs (new ideas) that fall in close proximity to known categories are rewarded more (considered more understandable) and outliers are rewarded less (more confusing). Finally, the amount of resonance a node produces decays over the lifetime of the node (becomes more boring), discouraging the agent from remaining in the highest density regions of its concept space. Thus, the intrinsic reward increases with residual change in the nodes' weights and the input's proximity to young, known categories, while it decreases with the growth of the network (relative entropy). This can be understood as an affinity for combinatorially creative ideas or changes to known ideas, while also enabling extensions to the concept space.

$$i = \frac{re_t}{e_{t-1}}|\mathbf{R}| \tag{2}$$

The exploration of the RL is modeled as a simple hypothesis evaluation, considering each possible next input and seeking the maximal resulting intrinsic reward (i). This approach is feasible in this limited domain, allowing the agent to look at pitches within a two octave range of the previously observed note. Duration, rhythm, dynamics, articulation and timbrel representations are not currently considered by the system. Thus at each decision point the agent need only evaluate at most 48 options, consider the reward resulting from each, and announce the choice of the maximally rewarding pitch.

The environment the RL works within is defined by the feature space it is presented with. We employ a feature vector of over one hundred elements derived from the pitch sequence (including pitch, pitch class, interval, interval class, change in interval, interval sign, register, average interval, minimum and maximum pitch over a given period, and direction). Due to the curiosity of the agent all of the available feature space will eventually (theoretically) be explored, producing every possible pitch sequence regardless of external aesthetic value. Human exploration is guided by many external factors (such as musical training, history, cultural conditions, etc.) which this agent lacks. Thus guiding the RL to make H-creative choices seems to be a function of extensive exposure (playing the corpus of recorded music for its evaluation) or the imposition of extrinsic reward measures (i.e. composing and constraining the feature space within which the agent may work), taking the place of a strict teacher. The latter seems like the more immediately viable solution but the former perhaps more true to human experience.

A. Limitations

As previously noted, this RL uses a highly restricted perceptual model that excludes many defining characteristics of music. As a result, the expressive space of this agent is very limited. Yet, given these constraints the material generated by the system exhibits novel developments, on a local scale. Exact repetitions are rarely seen but the musical texture is highly related, in pitch selection and interval patterns [9]. While the music can be interesting in small sections it clearly lacks any sense of macro direction, form, structure, or playing with a listener's expectations. This follows logically from the model set forth above, where the agent makes local choices, looking one note ahead to find a novel pattern.

Informing the RL's sense of musical direction may be accomplished with several different approaches. First, a hierarchical system of RL algorithms (as suggested by [7] and shown in [9]) would provide patterns and the same curious exploration at various levels, where each layer observes and guides the layer below. Thus the lowest layer (described previously) would make pitch choices, the next layer would make motivic choices, the next at a phrase level, etc. Each RL layer would watch the ART classifications of the lower layer and calculate intrinsic novelty rewards.

At the moment the feature encoding model treats all notes equally, regardless of emphasis in the musical texture. Human identification of melodic patterns is strongly informed by accentuation [10] and an appropriately analogous model may provide a significant improvement in the system's generation. Musical accents can be understood as an exaggeration or offset in a particular dimension, such as dynamics, register, rhythmic duration, metric placement, or contour (peaks and valleys, and approaching by a leap versus scalar motion). This extension may also lead to a ready incorporation of the additional musical elements that are currently being excluded from the RL's awareness (dynamics, rhythm, etc.).

Another approach is to develop a model of musical expectations that could parallel a human listener's experience. Knowledge of the operating style as well as general psychological principles inform a listener's expectations of how a musical work will unfold and the meeting or denying of these expectations is a proven informer of emotional experience [11]. Potentially, such a model could be developed for the RL to use, allowing the agent to formulate patterns to create musical direction that relate to a sense of expectation. Then the curious, novel exploration could intentionally fulfill or deny these expectations, and do so in the creative fashion argued thus far.

V. Conclusion

The RL design presented here suggests the potential for increased intelligence and creativity in automated music composition and improvisation. This in turn presents exciting possibilities for interactive work in which both the computer and the human agents are listening in a meaningful way and fundamentally changing in response to one another. However, the model employed betrays many limitations that may be overcome by a more general, creative RL model [2]. Functionally, the current design cannot discover algorithmic compression schemes, limiting its sense of novelty to feature pattern relationships, and its domain of exploration is highly constrained. Future work seeks to reveal further models and algorithms that can give the RL agent a way to produce historically meaningful music that can have value to a larger human audience.

Acknowledgment

The authors would like to acknowledge the support of eDream, the Institute for Advanced Computing Applications and Technologies, the Illinois Informatics Institute and the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign.

References

[1] J. McCormack, "Open problems in evolutionary music and art," *Applications of Evolutionary Computing*, pp. 428–436, 2005.
[2] J. Schmidhuber, "Driven by compression progress: A simple principle explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes," *Anticipatory Behavior in Adaptive Learning Systems*, pp. 48–76, 2009.
[3] M. Boden, *The creative mind: Myths and mechanisms*. Psychology Press, 2004.
[4] A. Newell, J. Shaw, H. Simon, and R. C. S. M. CALIF., *The processes of creative thinking.* Defense Technical Information Center, 1959.
[5] C. Davis and J. Bowers, "Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 32, no. 3, p. 535, 2006.
[6] G. Carpenter, S. Grossberg, and D. Rosen, "Fuzzy art: Fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural networks*, vol. 4, no. 6, pp. 759–771, 1991.
[7] R. Gjerdingen, "Categorization of musical patterns by self-organizing neuronlike networks," *Music Perception*, pp. 339–369, 1990.
[8] B. Smith and G. Garnett, "The self-supervising machine," in *Proc. New Interfaces for Musical Expression*, 2011.
[9] ——, "Reinforcement learning and the creative, automated music improviser," in *Proc. EvoMUSART*, April 2012.
[10] P. Pfordresher, "The role of melodic and rhythmic accents in musical structure," *Music Perception*, pp. 431–464, 2003.
[11] C. Krumhansl, "An exploratory study of musical emotions and psychophysiology." *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, vol. 51, no. 4, p. 336, 1997.