Assignment: Research Project
Course: Computational Creativity
Course Instructor: Rob Saunders

Students: Pieter Pierrot & Mauro Ricchiuti
Student IDs: 1089196 & 2462168

Date: December 11th, 2020

Title: The Drawing Apprentice Apprentice

Introduction

The initial and immediate interest of both authors lay with a co-creative system in the visual domain. The reason for this is twofold. First, working with visuals over other sensory modalities such as sound, is inherently easier and would give us more options in terms of skill and time. We did for a moment discuss making a system with which a user could play music in real time. To our knowledge, a lot of good commercial and open source creative support tools (CST) already exist in the music field, such as Jnana for Ableton Live [1] and mr Chainkov for VCVRack [10]. A Magenta Studio plug-in for Ableton Live [9], which makes use of Recurrent Neural Networks and Variational Autoencoders is available since 2019. Considering the weaknesses of RNN when it comes to real-time music generation [7], we thought about making a system intended to overcome these weaknesses. We abandoned the idea for the reason that it would be too risky and time consuming and could, in the end, defeat the intention of creating a fluid interaction. In addition, we agreed that there already existed a system that, from all the ones discussed in class and the ones we already knew of in the field of co-creativity, displays the most interesting results and underlying collaboration mechanics.

The Drawing Apprentice is a robust and influential system in the field of co-creativity, and its creator Nicholas Davis is a productive author in both theory and applications [2, 3, 4, 5, 8]. He and his team used methods and perspectives from cognitive science, art science, interaction design and computer science and AI to construct their CST. This approach yields a system that is easy to use by its design, and collaborates with the user in a way that both adds valuable contributions to the painting and inspires the user to be creative [2, 3].

The perspective Davis and his team take is a very user-oriented one. In the 2016 paper [3], they describe a method of analyzing whether or not the collaboration 'makes sense' to the user, invoking the cognitive scientific theory of *participatory sense-making* [6]. This framework was constructed to account for human interaction and collaboration; Davis and his team applied it to their computational system. To be able to make sense of the system, the user should be able to detect patterns, to a certain degree predict its contributions, or otherwise be able to understand the actions of the system. Interestingly, the understanding Davis and his team use for their system is that of "a

dynamical agentive systems point of view as an *interaction and coordination of two embodied agents* (italics from the original)". The embodied nature of this mutual process is important to Fuchs and The Jaeger, because the sense-making processes they envision stem from embodiment and its human imperatives: sociality, intentionality and movement.

This user-centered approach of The Drawing Apprentice is reflected in the voting mechanism, that functions as the main source of feedback to the system: the user can, after each turn, vote the contribution of the system either up or down (or don't vote). The system takes these votes into account and through them learns about the user's preference. This voting mechanism is crucial, as it defines the collaboration aspect of the system. One does not collaborate with a system or agent that cannot observe, react to, or learn from their own contributions. Put differently: The voting mechanism is what makes The Drawing Apprentice an interactive AI-powered tool, instead of a gimmick system (albeit a good one) that contributes to a drawing by pseudo-randomly selecting a piece of input and a transformation to execute. Davis and his team are aware of this fact. They describe, right before explaining the need for a smooth interface that helps lubricate the interaction [3]:

> "... Therefore, the most general design principle is whenever possible, offload any higher-level cognitive tasks to the user, i.e. enable the user to manually specify the boundaries within which the agent should operate throughout the interaction."

Further on in the text:

> "Achieving human-level collaboration is the ultimate goal for the type of co-creative system we designed. To do so, it is critical to understand the nature of participatory sense-making in the domain of artistic collaboration and the metrics users employ to evaluate their partner."

However, the voting mechanism is also the part of the system that caused the most confusion for the users. In the section on the results, the authors describe how the feedback mechanism was poorly understood and in some cases even omitted. Users reported frustration and uncertainty over how the voting affected the system. The overall positive responses on the interaction with the system reveal that the smart transformation and drawing implementation ensure a dynamic and intuitive interaction, but the lack of proper feedback to the system makes its claim to be a collaborative tool questionable.

This inspired us to try to implement a more rich and meaningful feedback mechanism into the overall interaction design of The Drawing Apprentice. The more direct the drawing specifics of the system depend on the user feedback, the more natural it will feel to the user. We believe that the confusion reported by the users of The Drawing Apprentice could be attributed to the nature of the Reinforcement Learning system. It needs a huge amount of training data before any real change can be perceived and such an amount cannot be obtained within the interaction of a single user. Gathering them from different users, on the other side, would still be hard and, moreover, meaningless for an individual user, as a drawing style is an intrinsically personal thing. We therefore replaced the RL approach with an evolutionary implementation that makes the feedback loop more direct. With this, we hope that any piece of feedback the user gives, will be immediately translated into a visible change in the systems output.

Implementation

The system was implemented fully in TouchDesigner, a visual programming language for real-time interaction and multimedia creation, which makes it convenient to develop both the visual and the interactive part. The core structure is similar to the one described in The Drawing Apprentice [3] and used in many CST's in the drawing domain: a paintbrush controlled by a simple UI and a generation system that can be collaborated with to create drawings. A user can draw with the paintbrush holding the left mouse button and a coloured line will be displayed. No object-recognition was implemented in this system (as was done in The Drawing Apprentice [4]); the expected interaction and collaboration with the system will take place on the level of lines and shapes.

Since we wanted to provide an alternative to the voting method and Q-learning used in The Drawing Apprentice, we choose to use an evolutionary approach for the generation of drawings. We chose to apply very simple transformations on the user input, with the aim of presenting the user with clean and intuitive contributions to their input on the screen. These transformations are evolved with the parameters as genotypes. The transformations are: scale and translate independently on X- and Y-axes, rotate, and distort (which applies noise to the input drawing coordinates) for a total of 6 parameters.

The interface is deliberately kept very basic. The user can change the diameter of the stroke using the "BrushSize" slider and its color using the appropriate UI widget. They can also set the color of the machine paintbrush, and change the size of the machine's stroke by moving the slider before it makes its draw. A "Reset" button is provided to completely delete the drawing from the canvas. Finally, the user can export the drawing that is currently on the screen with the export button. The image will be saved in the directory the TouchDesigner file is in.

The key component is of course the set of generation buttons. Each of the four generation buttons will trigger a different drawing based on your last recorded input. The input drawing is recorded as (x,y) locations in time, and a timer of 3 seconds will start when the left mouse button is released, after which the system considers the user interaction finished and will keep the recording buffered. This mechanism makes sure that the user is done taking their turn. The mechanism is implemented in the exact same way as in The Drawing Apprentice, and the 3 seconds seemed to be a good time interval for the turn-taking to feel natural. The user can draw multiple lines or shapes, but if they stop for 3 seconds, the recording will consist of everything since the last time they stopped for 3 seconds (or the start of the drawing).

By hovering the mouse above the generation buttons, the system visualizes different contributions on the screen. If the user likes one of these contributions, the particular generation button can be clicked and the machine draws it permanently and stores the parameters as a genotype for evolution. The different buttons create the following genotypes:

- Gen1 keeps the same genotype as chosen for the last generation and applies the same transformations to the input.
- Gen2 does a crossover, by randomly choosing parameters from the previous genotype.
- Gen3 behaves like Gen2, but uses a different random seed. Additionally, it adds a mutation, increasing every parameter with 10%.
- Gen4 generates a totally random set of parameters.

Buttons 1 through 3 are designed to be intuitive and direct, while button 4 is designed to create surprising and creative contributions. Of course, if the user likes none of the system's ideas, it can choose to keep on drawing and ignore the contributions for this specific line or shape.

Evaluation evaluation

For the evaluation of our system, we used 4 test persons that were given the instruction to make a drawing while collaborating with it. There was no intended goal, other than to make a nice drawing. Of these test persons, 2 were chosen because of their inherent affiliation with drawing, and the other 2 for the absence of such affiliation. It was told to all of the subjects that they were to collaboratively make a drawing with a computer system that aimed to complement and inspire them.

To get insight into the working of the system and its strengths and weaknesses (both its design and its potential for fruitful collaboration), we used a purely qualitative approach. This approach consisted of observation and transcription of the users' experience while collaboratively making the drawing, as well as a semi-structured interview afterwards (see Appendix I). The goal of the evaluation was threefold: to learn

- whether the users found it pleasant to work with the system
  → To test the interaction design, including the UI, and the value of the system as a drawing tool
- whether the users found the contributions suggested by the system were valuable
  → To test the creative potential of the system
- whether the users could make sense of the system's actions.
  → To test whether our improvement over the original Drawing Apprentice system

The responses on the usability and intuitiveness of the system were overall very positive. The UI was judged as being very clear and easy to use, and yielded no real surprises. One of the participants said[1]:

"The system was very clear! And I always get nervous from computer stuff, so using me as a participant is like starting with a huge disadvantage."

Such enthusiasm seemed to be shared among all participants. This can be interpreted as the fact that the minimum required quality of our system, an intuitive interaction, seems to have been met. One suggestion for improving the UI was made, namely the addition of a 'redo my last move'-button, that would give the user more control over what the drawing should look like. We did actually discuss this option before presenting the final version to the participants. However, it was not at all an easy feature to implement. Therefore, we decided to leave it out.

The participants were not very positive on the contributions of the system. Overall, the drawings suggested by the system were seen as random and incoherent, and participants reported they did not like the way the system drew off their initiative. For instance:

[1] These quotes are not literal quotes, but paraphrased transcriptions, as the user interviews were not audio-recorded.

"I think the system lacks inspiration. It is not like it comes up with stuff by itself, it just copies whatever I do."

Another participant made the remark:

"It seems like it has a lot of trouble following the lines that I drew."

Importantly, all participants made comments that lie along these lines, suggesting that the level of sophistication our system has might not be enough to produce valuable contributions to the human drawings. One participant observed a problem with the copying nature of our system. They said:

"Sometimes the system produced contributions that looked nice, but still it would be copying my unfinished drawings. The system's drawings would be valuable if it could finish my unfinished ideas."

This comment reveals that the lack of understanding the system has of what the user is drawing, semantically, keeps it from contributing in a valuable way. This is a problem, as copying the user's input and transforming it in some way is at the basis of the system's mechanics. However, working off the user's input (which does not have to mean copying, but also repeating, finishing, changing, extending, negating etc.) is also the main way in which the system reacts to the user, making it a collaborative CST. The original Drawing Apprentice system was extended with an object recognition module [4]. This might also be a necessary next step in the development of our system.

The main feature of our system, the evolutionary approach to drawing generation, with the different buttons producing different evolutionary results, was not entirely understood by the participants. After collaborating, some of them reported having little idea of the underlying differences between the buttons, let alone being able to exploit the differences and use them for more fruitful collaboration. For instance, one participant commented:

"I had some vague ideas about what could happen, but mostly playing with the generation buttons was random"

This feeling was echoed by another participant, who said:

"Sometimes it would re-create what I drew, but at other times the drawings seemed to be completely random"

Of course, it is possible that these participants used mostly button 4, which actually creates a random transformation. What is more likely, and fits better with the second report, as the drawings themselves were not random, but the transformations performed on the user's input was, is that it was just too hard to interpret and generalize over the rather nuanced differences the buttons provided. Let's say the user uses button 1 twice in a row (which means that the same transformation is applied). The first time after drawing a small blue line in the periphery of the drawing, the second time after drawing a big black circle in the middle. Even though the transformation is the same in both cases, it can be hard to figure it out for the user, as the contexts for the transformations are so vastly different.

One participant reported having a pretty good idea about what was happening. They said:

"I think collaborating went quite well. I felt like I had a pretty good idea of what was happening."

This more optimistic report is hopeful, but should not be taken at face-value. This user described how different buttons resulted in different translations, which was not exactly the mechanism that underlies the drawing generation. Even so, the fact that the user felt they had control over what was happening and could predict the outcome of the buttons is a good sign, as a good CST should be used with trust and opportunism to unlock its full potential. It is worth noting that this participant also was one of the non-drawers, who reported that they were happy to outsource some creativity to the system, to distract from the fact that they did not like drawing.

The system was considered by one of the participants with a lot of drawing affiliation as not appropriate for their professional practice, but apart from reporting of having fun playing with it, they also reported that the system helped her to produce something new:

"I drew things that I had never drew before, so in a way the system was inspiring me"

The same participant also tried the system with a drawing pad (WacomTablet), instead of a regular mouse(pad). They reported that this greatly increased the enjoyment of working with the system, as drawing with a pen on a surface is way more intuitive and akin to actual drawing than working with a mouse. This is an interesting observation, that should definitely be taken into account for anyone designing a CST for drawing[2].

Concluding remarks

Our system does not entirely fulfill its promises. The interaction design was successful, as our test users reported an intuitive UI and an easy-to-use system. This, however, was a basic requirement for success. More importantly, the evolutionary approach to drawing generation has not been developed far enough to call the overall system a grand success. This is no big surprise, as a fully developed co-creative drawing system is hard to produce in the time available to us for this project.

We believe, however, that the evolutionary approach has the potential of overcoming the problems faced in this study design. The way genotypes are represented through the parameters and the way the genotypes translate themselves to actual drawings based on the user input can be improved, made more intuitive or maybe explained deeply to the user before the start of a drawing session. This way, the link between buttons and drawings and the learning curve of the system will be more transparent. Of course, the object recognition module that was mentioned before will most likely be a big improvement to the system, as drawings do not consist of lines and shapes, but also of objects, and more complex structures. Finally, a shift of the target users towards kids, as suggested from one of the participants, could also give different meaning to the system: not as a CST per se, but as a fun tool for visual exploration.

[2] For a selection of collaborative drawings (with both the authors and the participants), see Appendix II

References

1. Colin Sullivan, 2012, Jnana, Max4Live plug-in, Retrieved from https://github.com/colinsullivan/Jnana
2. Davis, N. M. (2013, November). Human-computer co-creativity: Blending human and computational creativity. In *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*.
3. Davis, N., Hsiao, C. P., Singh, K. Y., Li, L., & Magerko, B. (2016, March). Empirically studying participatory sense-making in abstract drawing with a co-creative cognitive agent. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (pp. 196-207).
4. Davis, N. M., Hsiao, C. P., Singh, K. Y., & Magerko, B. (2016, September). Co-creative drawing agent with object recognition. In *Twelfth artificial intelligence and interactive digital entertainment conference*.
5. Davis, N., Hsiao, C. P., Singh, K. Y., Li, L., Moningi, S., & Magerko, B. (2015, June). Drawing apprentice: An enactive co-creative agent for artistic collaboration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition* (pp. 185-186).
6. Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the cognitive sciences*, *8*(4), 465-486.
7. Hadjeres, G., & Nielsen, F. (2020). Anticipation-RNN: Enforcing unary constraints in sequence generation, with application to interactive music generation. *Neural Computing and Applications*, *32*(4), 995-1005.
8. Karimi, P., Grace, K., Davis, N., & Maher, M. L. (2018, July). Creative sketching apprentice: Supporting conceptual shifts in sketch ideation. In *International Conference on-Design Computing and Cognition* (pp. 721-738). Springer, Cham.
9. Roberts, A., Engel, J., Mann, Y., Gillick, J., Kayacik, C., Nørly, S., ... & Eck, D. (2019). Magenta studio: Augmenting creativity with deep learning in ableton live.
10. Unless games, Mr. Chainkov, VCVRack plug-in, Retrieved from https://library.vcvrack.com/unless_modules/markov

Appendix I: Interview Questions

Note: These questions were intended as a guideline and inspiration to the interviewer. They could have been paraphrased in the interviews with the participants.

1. **Did you think the system was easy to use?**
   → Did the interface make sense
   → Was it easy to figure out how to use the system
   → Was it easy to make something nice

2. **Did you think the system was fun to use?**
   → Did you make something nice
   → Did you enjoy working with the system
   → Did you want to continue drawing

3. **Were the contributions the system made to the drawing valuable?**
   → Did they look good

4. **Did the feedback mechanism of the system make sense?**
   → Did you feel like you could predict what the system would draw
   → Was the difference between the different buttons clear

Appendix II: Drawings