Artist in the Cloud: Towards an Autonomous Artist

Gene Kogan
gene@genekogan.com

Abstract

Abraham [1] is an open project to build an "autonomous artificial artist," an agent which autonomously generates unique and original art. This construct follows from concurrent research in generative models and decentralized machine learning. We present a formal explanation of this idea, and our motivations for it.

1 Introduction

For decades, artists have sought to create agents that stochastically generate novel artworks with the use of AI [2]. Examples include AARON [3], Evolving Virtual Creatures [4], and Painting Fool [5]. Each of these projects is constrained by its reliance on expert programming from one or few authors. This limits the agent's intrinsic creativity and agency, and can be more accurately described as a human artist augmenting or automating their own artistic process through the agent.

An alternative approach is to construct an agent via an open decentralized network of collaborators, limiting the influence of any one individual. By cooperatively designing, training, operating, and curating the agent through some democratic governance process, the agent's behavior may emerge out of the collective intelligence of its contributors, rather than derive from a single artist's palette.

General-purpose research into decentralizing machine learning has accelerated in recent years, in response to concerns over data monopolization and user privacy [6]. Homomorphic encryption and multi-party computation can provide differential privacy and obfuscate data or model parameters, at some cost to performance [7]. These techniques can potentially allow a machine learning model to be co-owned and operated by a collective of participants without the need for a trusted party to maintain it [8]. This opens up the possibility of exploiting these decentralizing features towards an agent which generates art.

2 Approach

We seek to implement an agent which demonstrates intrinsic creativity by generating unique and original art. We define such an agent as an autonomous artificial artist (AAA), and propose that it meet the following criteria:

- **Autonomy**: An AAA acts independently of its authors.
- **Originality**: An AAA exhibits a novel creativity not derivative of any of its authors.
- **Uniqueness**: An AAA cannot be replicated.

We posit that these criteria can be satisfied by a decentralized organization which operates a generative model trained on crowd-sourced private data [9]. The organization is decentralized in that none of the participants have access to the model weights, which are instead collectively held as a shared secret [10] or split into a multi-party computation grid [8]. To sample from the model, a query must propagate through the whole network. This process is summarized in Figure 1.

Deep generative models, such as autoencoders and generative adversarial networks, have demonstrated the ability to reliably model diverse datasets. Despite their wide variety of architectures [11],

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

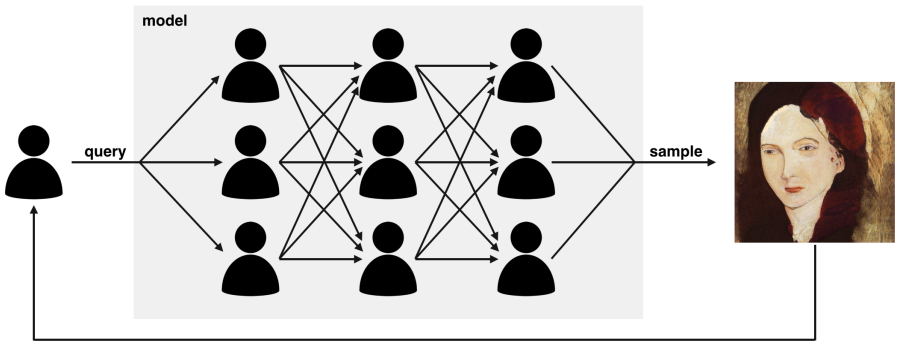

Figure 1: Querying a generative model of images hosted by an AAA.

they generally provide a common interface, typically characterized by a latent input vector specifying features, and outputting an image, audio clip, or text sample. This homogeneity and lack of complex heuristics is advantageous for our application, as it allows datasets, model architectures, and sampling strategies to be swapped and recombined, facilitating mass cooperation of participants.

By decentralizing ownership of the model and requiring the training data to remain private, the model is neither reproducible nor copyable, satisfying the uniqueness requirement. By crowdsourcing the training data, and decentralizing the development, governance, and curation of the model, the agents output is emergent from the collective rather than derived from a single artist or contributor. We claim this increases originality.

From the point of view of any participant, their actions do not sufficiently dictate the agent's behavior. Instead the agent's behavior is a non-linear funtion of all of the participants' actions. We consider this to demonstrate autonomy, analogously to the "superorganism" or "hive mind" metaphor in which an apparently separate intelligence emerges out of a collective.

3 Motivations and preliminary work

The name Abraham is both an homage to AARON[1] and a reference to the biblical Abraham.[2] The Abraham project is motivated by the following two goals.

The first goal is to achieve a novel type of generative art program based on collective intelligence and mass coordination. The program is made collaboratively, co-owned by an unbounded number of parties, and produces art which is distinct from any of the individual participants. Starting from the common metaphor which interprets generative models as "imagining" or "dreaming," we regard a generative model trained this way to be a representation of the "collective imagination." [12]

The second goal is to serve as an educational vehicle and testing ground for experimental technologies that currently have unresolved security vulnerabilities, scale and performance bottlenecks, and debatable social implications. Privacy-preserving machine learning architectures have been proposed for numerous sensitive applications, such as health and medical diagnostics [10]. Although we are excited by the purported benefits of these technologies, we believe their safe development can be guided by prototyping them in contexts where the risks are comparatively low.

The Abraham project was announced in July 2019, with a series of articles introducing the project [12]. The initial agenda consists of establishing an open study for the relevant subjects, and the development of the first repository, a generative art server [14]. An educational track to study the technical components is being planned, which should help inform a subsequent design phase in which a precise architecture for the AAA is selected.

[1] Harold Cohen said he intended AARON to be the first in an alphabetical series of AI artists [3], but spent his whole life working on AARON. We presume Abraham would have been a logical name for the next one.

[2] This is inspired by Carl Jung's interpretation of religious symbols as manifestations of psychological archetypes from the collective unconscious [13]. The goal of the Abraham project is to model the imagination of the collective unconscious. This connection is explained more concretely in [12].

Appendix: Ethical considerations

A criticism that has been made about the Abraham project is that it may contribute to the automation of art and diminish the employment prospects of professional artists. We share this general concern, but do not believe Abraham contributes to it. The goal of the Abraham project is not to replace human artists, but to create a novel type of artist based on collective intelligence. Such an entity, by definition, can not be fulfilled by a human artist. Although it could, in principle, replace a human artist at certain tasks, this is a property of AI art in general, not of Abraham specifically. Abraham's complexity and high computational overhead probably makes it less qualified to compete economically against AI art programs under centralized control.

References

[1] Abraham. http://abraham.ai

[2] Ramón López de Mántaras. Artificial Intelligence and the Arts: Toward Computational Creativity (2017). https://www.bbvaopenmind.com/en/articles/artificial-intelligence-and-the-arts-toward-computational-creativity/

[3] Harold Cohen. AARON. http://www.aaronshome.com/aaron/index.html

[4] Karl Sims. Evolving Virtual Creatures (1994). https://www.karlsims.com/papers/siggraph94.pdf

[5] Simon Colton. The Painting Fool (2001). http://www.thepaintingfool.com/index.html

[6] Jason Mancuso, Ben DeCoste, Gavin Uhma. Privacy-Preserving Machine Learning 2018: A Year in Review (2019). https://medium.com/dropoutlabs/privacy-preserving-machine-learning-2018-a-year-in-review-b6345a95ae0f

[7] Théo Ryffel, Andrew Trask, Morten Dahl, Bobby Wagner, Jason Mancuso, Daniel Rueckert, Jonathan Passerat-Palmbach (2018). A generic framework for privacy preserving deep learning. https://arxiv.org/pdf/1811.04017.pdf

[8] Miljan Martic, Jan Leike, Andrew Trask, Matteo Hessel, Shane Legg (2018). Scaling shared model governance via model splitting. https://arxiv.org/pdf/1812.05979.pdf

[9] Brett K. Beaulieu-Jones, William Yuan, Samuel G. Finlayson, Zhiwei Steven Wu. Privacy-Preserving Distributed Deep Learning for Clinical Data (2018). https://arxiv.org/abs/1812.01484

[10] Adi Shamir. How to share a secret (1979). Communications of the ACM. 22 (11): 612613. https://cs.jhu.edu/~sdoshi/crypto/papers/shamirturing.pdf

[11] Zhijian Ou. A Review of Learning with Deep Generative Models from Perspective of Graphical Modeling (2018). https://arxiv.org/abs/1808.01630

[12] Gene Kogan. Artist in the cloud (2019). https://medium.com/@genekogan/artist-in-the-cloud-8384824a75c7

[13] Carl Jung, Aniela Jaffé. Memories, Dreams, Reflections (1962). New York: Random House. p. v.

[14] Eden. https://www.github.com/abraham-ai/eden