ACADEMIA | Letters

Are Agent-based Models Universal Approximators?

Joseph Shaheen

Universal approximation functions are well known and studied in canonical mathematics. Here we theorize the existence of an independent class of universal approximation agents and agent-based models. We draw upon historical references from mathematical analysis, the development of machine learning and the agent-based modeling lines of inquiry.

1 Introduction

Approximation theory and real analysis offer a class of theorems and proofs dubbed the universal approximations theorems. This collection of pure mathematics provides powerful mechanics and algorithms for approximating output datasets and their underlying generative functions. Three well-known examples are the polynomial functions (equation 1), the architecture of a convolution neural network and feed-forward neural networks.

$$f(x) = x^7 + ax^3 + bx^2 + cx + 1 = 0 \tag{1}$$

The origins of universal approximation theory can be traced back to the Second International Congress of Mathematicians in Paris, France in the twilight of the 20th century. In this meeting of venerated mathematicians,

David Hilbert posited 10 of 23 difficult problems important to the progress of mathematics. One commonly known–was Hilbert's 13th problem; He considered whether the solution of a polynomial of degree 7 (Equation 1) can be re-written as the composition of functions

Corresponding Author: Joseph Shaheen, jshaheen@gmu.edu
Citation: Shaheen, J. (2021). Are Agent-based Models Universal Approximators? *Academia Letters*, Article 1142. https://doi.org/10.20935/AL1142.

with only 2 variables. Herein, we could consider the variables *X* (*a, b, c*) as the solution function to be decomposed, and a function to be a mapping of one or more inputs to one or more outputs.

In the 1950s Andrey Kolmogorov and Vladimir Arnold solved this problem, first for continuous functions with three variables and then for two variables (Arnol'd, 2009, pp. 5–9), and shortly thereafter for one variable (Arnol'd, 2009, pp. 25–47). These important proofs show that any continuous function $f:\ [0,\ 1]n \to \mathrm{R}$ can be decomposed to a combination of inner and outer functions of a single variable (Equation 2).

$$f(x_1, ..., x_n) = \sum_{j=0}^{2n} \phi_j \left(\sum_{i=1}^{n} \psi_{i,j}(x_i) \right). \tag{2}$$

Further developments to what was then dubbed the Kolmogorov-Arnold Representation Theorem (KART) took place in the next three decades. However, Kolmogorov and Arnold's proof of existence did not offer a method by which an n-dimensional function can be decomposed into a one-dimensional function constructively. It only *proved* that it was possible. This remained the case until prominent scholars in neurocomputing–today's deep learning–recognized the linkage.

The first such recognition came from Robert Hecht-Nielsen (1987). Hecht Nielsen correctly argued that even though applied mathematics had failed to "[find] a significant use for [KART]" that this was "*not* the case in neurocomputing". Hecht-Nielsen (1987) also recognized that at the time of his publication no constructive method to impute the inner functions were known, but he predicted their arrival. As we know now, this was a prescient prediction. Over the next few decades, a concerted effort to develop the necessary methods tasked with finding inner functions (the processing layers) in the study of neural networks had begun and more theorizing on the linkage between universal approximation and *neurocomputing* (Hornik et al., 1989; K˚urkov´a, 1992; Kosko, 1994; Chen and Chen, 1995) was completed. Concurrent with the development of the mathematical strength to discover the core of universal approximators, developments in computing technologies over the next two decades left little doubt so as the usefulness of universal approximation.

Corresponding Author: Joseph Shaheen, jshaheen@gmu.edu
Citation: Shaheen, J. (2021). Are Agent-based Models Universal Approximators? *Academia Letters*, Article 1142. https://doi.org/10.20935/AL1142.

2 Agent-based Modeling

Running concurrently with the development in neurocomputing and capitalizing on the growth in computing technologies was the agent-based modeling and simulation frameworks of the 1990s.

The birth of agent-based models is commonly associated with Schelling's segregation model (Schelling, 1971) first conceived using a roll of pennies and quarters (agents) and a large sheet of graphing paper (the board)–a humble beginning–but an impactful one. Agent-based modeling allows for the modeling of qualitative interactions based on difficult-to-mathematically-define rule-sets ultimately describing some bottom-up interactive process. Nevertheless, the process of agent-based modeling describes a mapping between one or more inputs and one or more outputs often stationed on the real number line. By classical mathematical definitions, this is still a mapping and consequently–a function–albeit a procedural one. Agent-based models permit the use of heterogeneous mappings for each individual agent in a given system making them particularly useful for the study of complex systems (Axtell, 2000). The method has deep roots in complexity science fielding sub-domains such as network science, chaos theory, and evolutionary computing (Mitchell, 2009).

Schelling's model works by assigning agents a "happiness" score based on the number of (other) agents in their spatial vicinity who belong to their own group (e.g. black/white, red/blue, male/female). Agents can move randomly across a two-dimensional board and this movement is subordinate to their level of happiness. If each individual agent finds that they are unhappy with their current positioning then they move in the next iteration of the model run. If not then they remain in their current placement. The player of the model can set the parameter value for happiness under which an agent would be willing to move or remain in their position. This parameter can be distinctly applied to each agent or as a global property of the model. Thus, Schelling's model produces a mapping from each agent's threshold for happiness–an attribute set locally or globally–to an agent's positioning and final happiness score. Consequently, this is also a mapping to any macro property of the system (the board) we wish to investigate. Schelling's efforts focused on segregation and the clustering of communities.

Historically, such a model has proven difficult for analytical mathematical analysis and in no uncertain terms impossible to solve when the evaluation mechanism becomes multivariate, multi-variable, or multidimensional.

Brandt et al. (2012) embarked on a valiant effort to analytically solve the one-dimension case under several strong assumptions, but considered the two-dimension model (Figure 1) to be an "open problem". Thus, *in silico* computation remains the defacto method of solving

Corresponding Author: Joseph Shaheen, jshaheen@gmu.edu
Citation: Shaheen, J. (2021). Are Agent-based Models Universal Approximators? *Academia Letters*, Article 1142. https://doi.org/10.20935/AL1142.

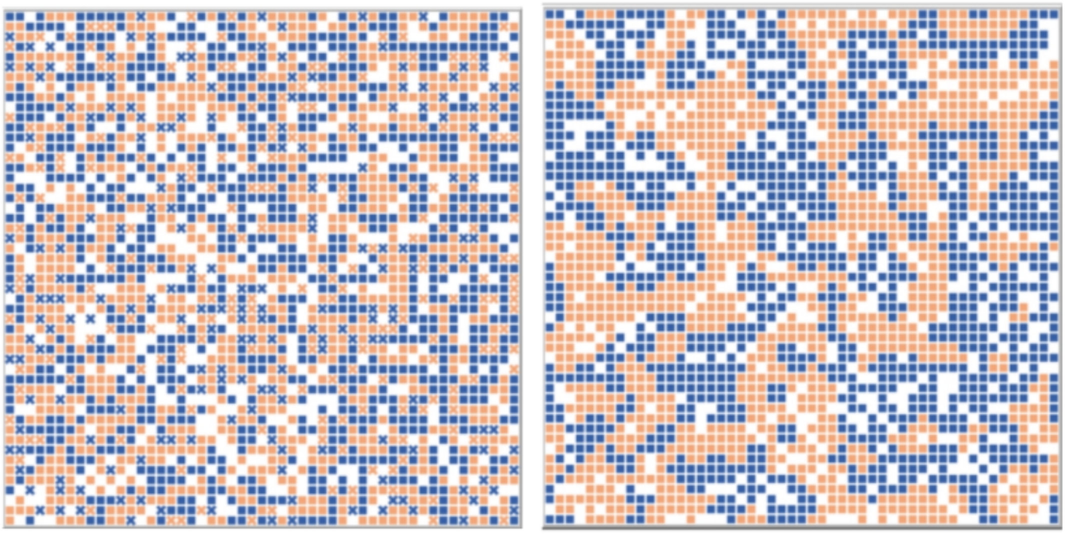

Figure 1: Schelling's segregation model in NETLOGO at instantiation (a) and after simulation run (b)with two groups (blue and orange)

this system–specifically– computation through *simulation*.

While Schelling's model was the first, his work on this front was not recognized until the mid-1990s in spite of his Nobel prize-winning efforts on collective dynamics and conflict resolution. The recognition came with the agent-based model *SugarScape* of Epstein and Axtell which crystallized the notion of many agent interactions producing system-level properties (Epstein and Axtell, 1996) or as Schelling had previously christened it in his seminal book "Micromotives and Macrobehaviors" (Schelling, 1978).

Sugarscape was not the first work to attempt to re-discover the ethos of connecting heterogeneous agents with system-level properties through simulations e.g. Batty (1976); Huberman (1989); Carley (1991); Albin and Foley (1992); Axelrod (1993); Resnick (1994). However, it represented an early success in the use of increasingly available computing power and an interest-generating manuscript. Epstein et. al. used–by even contemporary standards–ubiquitous programming technologies to do what Schelling did with pennies, quarters and graphing paper. They modeled a small society comprised of many heterogeneous agents on a two-dimensional landscape similar to Schelling's model.

Sugarscape modeled society as agents on a two-dimensional plane that sought to maximize their 'sugar' intake-analogizing materiel resources (Figure 2). Agents would scour the plane and consume sugar randomly placed at the instantiation of the simulation. When they metabolized the sugar at heterogeneous rates, they would then move randomly to a new area

Corresponding Author: Joseph Shaheen, jshaheen@gmu.edu
Citation: Shaheen, J. (2021). Are Agent-based Models Universal Approximators? *Academia Letters*, Article 1142. https://doi.org/10.20935/AL1142.

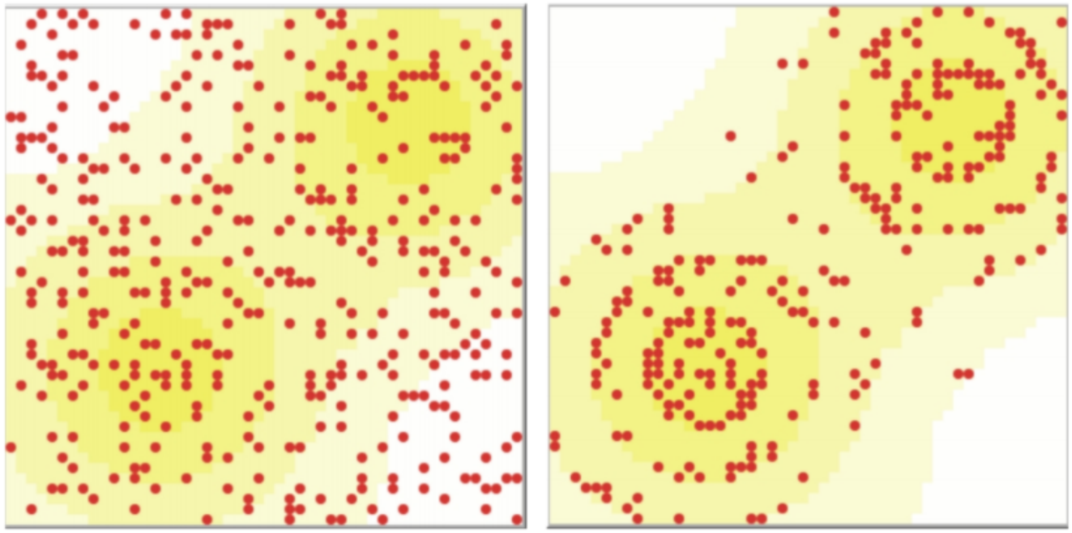

Figure 2: Sugarscape in NETLOGO at instantiation (a) and after simulation run (b). Agents are in red dots and sugar distribution represented by the vibrancy of yellow patches.

where sugar was yet to be consumed. Like Schelling's model, macro-level properties emerged from this rule-set. The authors tested variations of this model purported to produce patterns often found in real human systems. They argued that this simple model, and variants of it, produced well-known real-world societal-level properties such as wealth inequality, migration, social organization, cultural dissemination, trade dynamics, disease transmission, and even warfare. They made no argument that the list of all human-centric relational processes can or may be produced by their simple model but offered qualitative evidence that the model captures–or in more convenient terms–*approximates* reality quite well.

The departure from Schelling occurred in that Epstein and Axtell's agents were given a different set attributes to evaluate (inputs) using differing behavioral rules (processors) and the authors focused their attention on different outcomes (outputs)–a different mapping from inputs to outputs, but a mapping nonetheless. What no variation of Epstein et. al's Sugarscape produced is spatial segregation as an aggregate outcome, even when considering that it shares five of the more salient model design properties with Schelling's model: 1) agents are placed on a two-dimensional grid and cannot overlap spatially 2) agents evaluate whether there is space to move to another location before moving there, 3) agents evaluate their immediate surroundings but make no evaluation of the global properties of the system and, 4) agents move to a random position at every turn of the simulation run given some condition and, 5) space is limited.

Corresponding Author: Joseph Shaheen, jshaheen@gmu.edu
Citation: Shaheen, J. (2021). Are Agent-based Models Universal Approximators? *Academia Letters*, Article 1142. https://doi.org/10.20935/AL1142.

One may consider that *Sugarscape* and Schelling's model have little else in common. Yet, one of the more prominent conclusions inferred by both models is the production of inequality in sugar distribution on the one hand and segregation on the other. Notably, Schelling's model produces inequality in "happiness" but this is often not the focus of scholars' interest in the model. On the other hand, we can say that *Sugarscape* produces segregation in sugar wealth but this is not a common way of describing wealth inequality. In other words, whether it is inequality or segregation both models produce partitioning and inequity between the agents, given standard model run parameters and mechanics. Under the right parameter choices and aligning the modeler's viewpoint to measure the output which is directly related to the input (happiness to happiness distribution, sugar to sugar distribution), both models produce a similar conclusion: a skewed distribution of the input in question.

It is in this finding that the convergence of the mapping function be tween the segregation model and *Sugarscape* can be extracted. The agents in *Sugarscape* compete for sugar wealth through spatial movement so it is only natural to expect that inequity will occur in relation to the quantity that is contested i.e. sugar. Similarly, agents in Schelling's model compete for happiness through spatial movement, and so inequity occurs in relation to the contested quantity as well i.e. happiness. Consequently, (happiness) segregation occurs as a manifestation of the contested space. Since both con tested quantities are based on space, with sugar only accumulated in space where it is available and happiness only increased when space is available to move close to in-group members, one may conclude it is then space that is contested. Sugar and happiness are simply proxies for space in the models. This similarity, we claim, is an element of universality.

3 A Question

Combining what we learned from the purely mathematical turned machine learning line of inquiry with what we can qualitatively deduce from our basic comparison of two important agent-based models, we find sufficient evidence to simply pose the question: Since Kolmogorov (1957) guarantees the existence of single-variable approximators to any multivariate function, and since agents represent a mapping between one or more inputs to one or more outputs–a multivariate function, can we conclude that universal approximating agents and universal approximating agent-based models exist?

This question only generates more questions but we believe that the answers could generate a golden age in modeling theory.

Corresponding Author: Joseph Shaheen, jshaheen@gmu.edu
Citation: Shaheen, J. (2021). Are Agent-based Models Universal Approximators? *Academia Letters*, Article 1142. https://doi.org/10.20935/AL1142.

4 Acknowledgments

This research was supported by an appointment to the Intelligence Community Postdoctoral Research Fellowship Program at George Mason University, administered by Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Office of the Director of National Intelligence.

References

Albin, P. and Foley, D. K. (1992). Decentralized, dispersed exchange without an auctioneer. A simulation study. *Journal of Economic Behavior and Organization*, 18(1):27–51.

Arnol'd, V. I. (2009). *Collected works: Representations of Functions, Celestial Mechanics and KAM Theory, 1957–1965*. Springer, London; New York, 1st ed edition.

Axelrod, R. M. (1993). A Model of the Emergence of New Political Actors. Technical report, Santa Fe Institute.

Axtell, R. (2000). Why agents? On the varied motivations for agent computing in the social sciences. *Center on Social and Economics Dynamics - The Brookings Institution*, (17):1–23.

Batty, M. (1976). *Urban Modelling. Algorithms, Calibrations, Predictions*. Cambridge Urban and Architectural Studies. Cambridge, MA.

Brandt, C., Immorlica, N., Kamath, G., and Kleinberg, R. (2012). An analysis of one-dimensional Schelling segregation. In *Proceedings of the Annual ACM Symposium on Theory of Computing*, pages 789–803.

Carley, K. (1991). A Theory of Group Stability. *American Sociological Review*, 56(3):331.

Chen, T. and Chen, H. (1995). Universal Approximation to Nonlinear Operators by Neural Networks with Arbitrary Activation Functions and its Application to Dynamical Systems. *IEEE Transactions on Neural Networks*, 6(4):911–917.

Epstein, J. M. and Axtell, R. (1996). *Growing Artificial Societies: Social Science From the Bottom Up*. The Brookings Institution.

Hecht-Nielsen, R. (1987). Kolmogorov's Mapping Neural Network Existence Theorem. In *Proceedings of the International Conference on Neural Networks*, volume 3, pages 11–14.

Corresponding Author: Joseph Shaheen, jshaheen@gmu.edu
Citation: Shaheen, J. (2021). Are Agent-based Models Universal Approximators? *Academia Letters*, Article 1142. https://doi.org/10.20935/AL1142.

IEEE Press New York.

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366.

Huberman, B. (1989). The ecology of computation. In *Digest of Papers. COMPCON Spring 89. Thirty-Fourth IEEE Computer Society International Conference: Intellectual Leverage*, page 362. IEEE Comput. Soc. Press.

Kolmogorov, A. N. (1957). On the representation of continuous functions of several variables as superpositions of continuous functions of one variable and addition. *Dokl. Akad. Nauk SSSR*, 114(5):953–956.

Kosko, B. (1994). Fuzzy systems as universal approximators. *IEEE Transactions on Computers*, 43(11):1329–1333.

K˚urkov´a, V. (1992). Kolmogorov's theorem and multilayer neural networks. *Neural Networks*, 5(3):501–506.

Mitchell, M. (2009). *Complexity: A Guided Tour*. Oxford University Press. 8

Resnick, M. (1994). *Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds*. MIT Press.

Schelling, T. C. (1971). Dynamic Models of Segregation. *The Journal of Mathematical Sociology*, 1(2):143–186.

Schelling, T. C. (1978). *Micromotives and Macrobehavior*, volume 2. W. W. Norton & Company.

Corresponding Author: Joseph Shaheen, jshaheen@gmu.edu
Citation: Shaheen, J. (2021). Are Agent-based Models Universal Approximators? *Academia Letters*, Article 1142. https://doi.org/10.20935/AL1142.