Statement of Professional Goals and Objectives

Garrett E. Katz
April 1, 2024
http://web.ecs.syr.edu/~gkatz01

This document lays out my professional aspirations and surrounding context in chronological fashion, starting with my academic preparation and up through my current and future goals.

I. Origins

My academic journey has traced a winding and humbling path, from philosophy through math to biology and then computing. As an undergraduate, I was drawn to philosophy of mind; and especially interested in the nature of consciousness and the limits of human knowledge. I became fixated on the idea underpinning philosophical skepticism: that nothing can ever be known with absolute certainty, except perhaps logically necessary "truths" – not fully appreciating at the time that even logic itself is an evolving collection of mortal ideas.

Retreating inward to the realm of abstract logical thought, I went on to pursue a graduate degree in mathematics. I learned much mathematics in those years, which has since served me well, but perhaps the most important lesson was the severe limit of my own intelligence. I regularly faced problems I was unable to solve without weeks of grueling effort, and what felt like rare strokes of luck, if at all. It became apparent that meaningfully expanding the frontier of mathematical knowledge is an exceedingly difficult prospect, dependent on sparks of insight that are few and far between for most, and certainly for the likes of me.

There were two other important bends in the road at this stage of my academic journey. The first was a shift towards computer science in my elective coursework. I found that I had a certain aptitude for algorithmic thinking, and thoroughly enjoyed it. Computer programs are much like formal proofs; in some sense more tangible but no less rigorous.

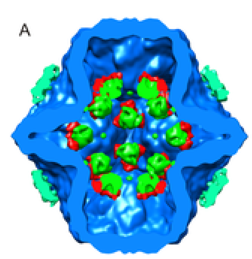

Figure 1: A 3D reconstruction we computed of Cystovirus $\phi6$.

The second was an opportunity to work as a research technician in a biology laboratory on campus. In this role I was using computational methods, such as expectation-maximization and Fourier transforms, to infer the 3D structure of biological virus particles from their 2D cryo-electron microscopy images. This experience gave me a baseline familiarity with biology and my first experience conducting research and publishing results [1, 2]. It also forced me out of my comfort zone of logical ideals to confront the noisy and messy nature of concrete, real-world problems. Most importantly, it exposed me to sophisticated computational techniques, and showed me the pivotal role that computing can play in the process of scientific discovery.

At this point, my sense of scientific purpose began to converge. I decided to contribute to the science of reasoning itself – understanding how it emerges and can be cultivated in biological brains, but also how it can be reversed-engineered in machines. My ultimate goal was to help humankind transcend its current limits in domains of scientific discovery and mathematical knowledge. I moved to Maryland where I had the incredible fortune of pursuing a doctorate under James Reggia, a multidisciplinary expert with a Ph.D. in computer science and M.D. in clinical neurology.

II. Doctorate

My doctoral work explored various aspects of automated reasoning in the context of humanoid robotics. The focus was on robotic imitation learning, in which a human operator "programs" a robot by demonstrating how to perform a task, rather than writing code. This enables robotics applications in which the human end-user may not have expertise in computer science, such as assisted living or distaster recovery.

Figure 2: The Baxter robot, our hardware platform for robotic imitation learning.

Much previous work on robotic imitation learning targeted the "motor level," meaning that the robot imitated the precise movements of the demonstrator, such as the trajectory of the hand. This is a non-trivial problem in its own right, due to differences between human and robot embodiments, among other things. However, motor-level imitation learning does not generalize well when task details change, such as the positions of relevant objects in the environment.

My contribution was a "cognitive-level" imitation learning approach [3], in which the robot does not imitate the actions of the human demonstrator, but rather their goals, such as the final states or arrangements of relevant objects. This approach involves two forms of automated reasoning. First, cause-effect reasoning is used to infer the likely goals of the demonstrator, based on their observed actions. Second, when task details change, automated planning is used to accomplish the same goals with potentially different motor-level actions that are more appropriate to the new situation.

In the parlance of artificial intelligence (AI), this approach was implemented "symbolically," meaning that it used traditional algorithms, data structures, and logical formalisms to represent and process objects, states, actions, and goals. This is in contrast with prevailing deep learning architectures today, which are called "sub-symbolic" because their representations and control laws are implicit in their myriad numerical parameters. However, a major issue in sub-symbolic models is their relative lack of interpretability. This is of particular concern in application areas involving human-machine teaming, where the human's wellbeing could be jeopardized. Intepretability is also important in other settings related to abstract reasoning. For example, humans are more likely to accept computer-assisted proofs if the proofs are simple and interpretable.

My doctoral robotics work laid the foundation for my current professional goals in several ways. First, it prepared me to study reasoning that is *grounded* in one's physical embodiment. Even the most abstract mathematical concepts do not form in the vaccuum of platonic ideals. Rather, a mathematician's ingenuity often stems from visual, spatial, and physical intuitions they have developed over a lifetime of sensorimotor experience and development. Second, it familiarized me with research involving human participants, since this research involved a substantial human-robot interaction component [4]. Lastly, it positioned me well to focus on the issue of interpretability, owing to the primarily symbolic approach [5].

III. Descent into the deep

Near the end of my dissertation work and in my subsequent postdoctoral appointment, I began investigating interpretability in sub-symbolic systems as well. These investigations involved two main research threads. The first thread was concerned with interpreting the complex dynamics of recurrent neural networks. Extant literature had shown that trained recurrent neural networks often possess meta-stable fixed points, which govern discrete steps and branch points in their observed behavior. The learned behavior can therefore be interpreted by identifying and linearizing dynamics around these fixed points. However, identifying fixed points is equivalent to solving non-linear systems of equations, and as such constitutes a challenging global search problem. I discovered a new technique for solving non-linear systems of equations, based on numerical continuation of novel mathematical objects I dubbed "directional fibers" [6]. This technique is useful in interpreting sub-symbolic AI, but also more generally for non-convex optimization, since the stationary points of an objective function are fixed points of its gradient field.

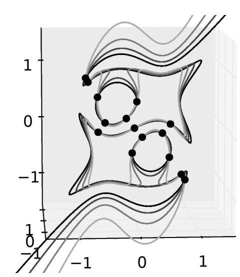

Figure 3: Numerical continuation of directional fibers.

The second thread focused on neurosymbolic integration. I developed a system that "compiles" symbolic, human-authored knowledge into a neurocomputational representation [7]. The system, which I called the "neural virtual machine (NVM)," is an architecture that – like natural intelligence – can faithfully emulate symbolic procedural knowledge using a purely neural implementation. This system forms a basis for interpretable sub-symbolic systems, since their behavior can be programmed by a human rather than learned from data. However, the neural representation is also amenable to data-driven fine-tuning when desired.

IV. Framework

My long-term professional goal has not changed since I began my Ph.D.: I still aim to make meaningful contributions to the science of reasoning, including automated reasoning in machines as well as reasoning processes in humans – especially students. Over the past five years I have worked towards this goal within a framework of "vertically-integrated" intelligence, by which I mean integration across all levels of intelligence in a unified system, from embodied low-level sensorimotor control to high-level abstract reasoning. One advantage of my winding academic journey is that it prepared me with some degree of interdisciplinary fluency across this "full stack" of intelligence, including biology and neuroscience, numerical optimization and control, sub-symbolic methods, automated logical reasoning, and human behavioral science. Since joining Syracuse University, the research threads initiated in my doctoral and postdoctoral work have been substantially expanded and integrated, as detailed in the following sections.

Sensorimotor Robotic Learning and Control

Three of my doctoral students and I have focused respectively on perception, locomotion, and manipulation in the full-body Poppy humanoid and the 6-degree-of-freedom Poppy Ergo grippers. These are all very challenging low-level problems and prerequisites for any robotics system, even when the long-term goal is to study embodied abstract reasoning. For the perception component, we have developed a predictive model that can anticipate robotic falls before they occur, based on the robot's egocentric visual input stream [8]. For the manipulation component, we have built a grasp planner on top of the NVM [9]. As such it can follow human-authored block-stacking algorithms, but also improve upon those algorithms with additional experience-driven reinforcement learning. We also have two papers under review on robot locomotion with robust reinforcement learning [10] and symmetric-gait reward shaping [11]. We have also presented an invited workshop paper on preliminary work combining all three sub-systems [12].

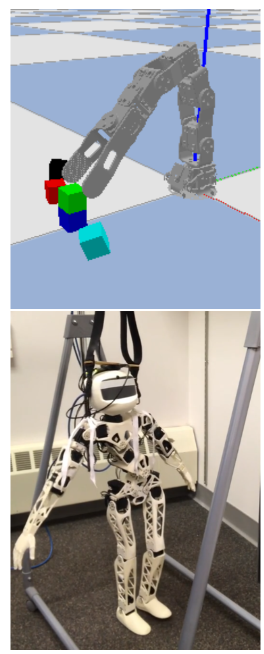

Figure 4: A simulated Poppy gripper stacking blocks and real Poppy humanoid taking a step.

Over the next five years, the primary objective is to equip the Poppy humanoid with gripper hardware and develop a robust whole body controller that can navigate and manipulate its environment. This will lay the foundation for studying emergent reasoning and planning capabilities, grounded in an embodied, NVM-based neurosymbolic controller.

Sub-symbolic Learning and Numerical Methods

My lab has cultivated a core skillset in (deep) sub-symbolic learning through our own robotics applications [8] and collaborations with others [13,14]. In addition to these published works, I meet regularly with a bioinformatics research group and neuroscience research group to explore impactful applications of deep learning and other AI methodologies. In the coming years I aim to publish papers and secure grant funding with members of both groups to broaden the impact of my lab's work.

I have also continued developing and applying the directional fiber methodology for optimization and dynamical system analysis. This includes work on global optimization [15], analysis of novel associative learning rules [16], and reverse engineering of reservoir computers trained on mathematical weather models [17]. One of my doctoral students has also found other ways to use numerical continuation in a deep learning setting [18]. His previous and ongoing work has shown it is possible to maintain a near-constant loss in a primary training objective while optimizing secondary objectives, such as smoothness, robustness, or sparsity of the network. In other words, he has developed a multi-objective optimization method based on numerical continuation. Scaling numerical continuation methods (which are often second-order) to very large models and datasets is a non-trivial research problem in its own right, and the topic of his upcoming dissertation defense.

One major drawback of deep learning is its exorbitant computational cost. In addition to negative environmental impacts, this creates a significant bottleneck for members of my lab and many others that lack the high-performance computing resources available to industry. In response I have also initiated research into single-pass learning that is more computationally expedient. Single-pass learning also has an added synergy with my work on the NVM, because the NVM uses single-pass learning rules to emulate random-access memory. Several of my students and I recently submitted a paper where we proved a theoretical hardness result about a large family of single-pass learning rules and their data-fitting capacity [19]. In the coming years, I aim to further broaden the scope of these hardness results, and also "close the optimality gap," by designing specific single-pass learning rules that attain their theoretical limits. This would have broad impact for the single-pass learning research community and also improve the scalability of the NVM.

Neurosymbolic Automated Reasoning

The NVM has served as a core foundation for much of my research since my appointment began, and my planned research in the years ahead. With collaborators I have used the NVM to build robotic controllers [9], model neurological cognitive disorders [20], solve program induction tasks [21], emulate LISP implementations of unification and other algorithms [22], perform abductive inference [23], and probe questions about artificial consciousness [24]. Going forward, I aim to make the NVM substantially more scalable than its current form by improving its underlying single-pass learning algorithms, as mentioned above.

My research on automated reasoning is not limited to the NVM. I have published work at AAAI on the abstract reasoning task of automated algorithm design [25], and in the coming

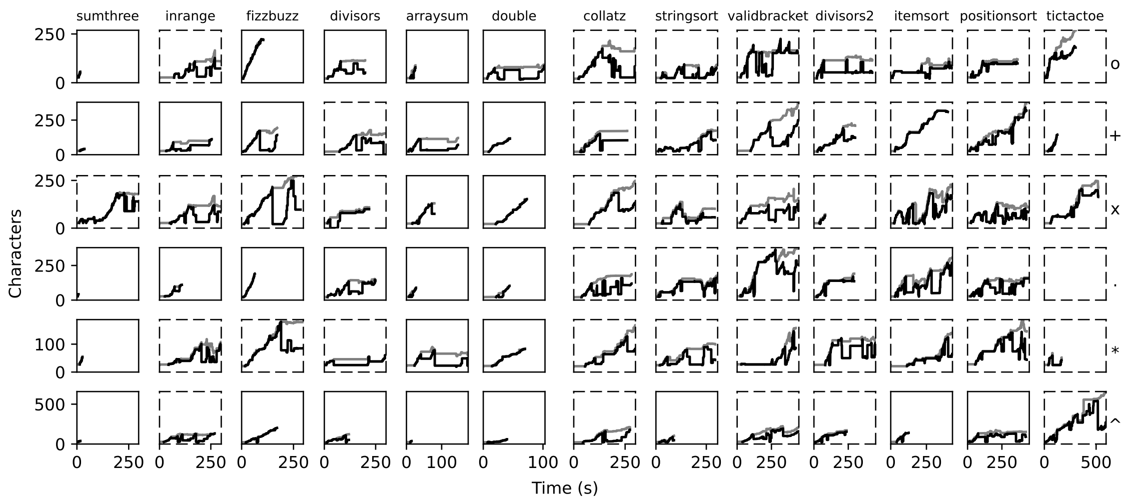

Figure 5: Cursor dynamics during student problem solving.

years I aim to publish work on automated mathematical theorem proving. I have brought myself up to speed on recent literature in this area by organizing a workshop at AAAI 2023, preparing my recent topic course on the subject, and developing a machine-learning-friendly Python interface to Metamath (a major interactive theorem prover used in this area).[1]

Human Problem Solving

In recent years I have continued collaborating with colleagues at University of Maryland's kinesiology department who analyze human cognitive-motor performance in a human-robot teaming context [4,26]. This work studies both performance and mental effort of human subjects, with and without robotic teammates, on tasks involving both cognitive-level problem solving and low-level motor-control (such as Towers of Hanoi and the Rush Hour game).

I have also initiated my own human subject research into the problem-solving process of computer science students, as they attempt programming exercises. This research is intended to inform my work on human-like automated reasoning, but also to elucidate the reasoning process of computer science students so we can ultimately improve their learning outcomes. In an IRB-approved pilot study, I keylogged several Syracuse students during coding exercises and correlated their cursor dynamics with their performance on an automated test suite. Figure 5 shows the cursor position over time for 6 students on 13 problems, exhibiting highly non-linear movement particularly in attempts where they were struggling. In this pilot I also observed a highly bimodal distribution in performance: Students tended to pass either all or none of the automated tests in a given attempt. This raises questions such as the extent to which students are problem-solving as opposed to recalling memorized solution strategies, and warrants further study. This work was done with the assistance of a summer REU student, and a manuscript on our results is under review [27].

[1] https://github.com/garrettkatz/mmpy

V. The long game

Over the long-term, I hope to elucidate and harness the power of *creativity* in abstract reasoning and problem solving. Famous theoretical results, such as the infinitude of the primes or uncountability of the reals, have proofs that are remarkably simple and interpretable in hindsight – but were originally discovered with extraordinary creativity. At the more tangible level of computer science education, an aptitude for creative problem solving is arguably just as important as an attention to rigor and formal rules.

However, the source of creativity in humans remains mysterious, and most automated reasoning systems, whether neural or symbolic, do not solve problems in the same creative ways that humans do. Deeper insight into the computational principles behind creative problem-solving can help us improve the human reasoning process, build more interpretable AI systems, and reach new heights in the endless pursuit of knowledge.

References

[1] G. Katz, H. Wei, A. Alimova, A. Katz, D. G. Morgan, and P. Gottlieb, "Protein p7 of the cystovirus φ6 is located at the three-fold axis of the unexpanded procapsid," 2012.

[2] G. Katz, Y. Benkarroum, H. Wei, W. J. Rice, D. Bucher, A. Alimova, A. Katz, J. Klukowska, G. T. Herman, and P. Gottlieb, "Morphology of influenza b/lee/40 determined by cryo-electron microscopy," *PloS one*, vol. 9, no. 2, p. e88288, 2014.

[3] G. Katz, D.-W. Huang, T. Hauge, R. Gentili, and J. Reggia, "A novel parsimonious cause-effect reasoning algorithm for robot imitation and plan recognition," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 2, pp. 177–193, 2018.

[4] T. C. Hauge, G. E. Katz, G. P. Davis, K. J. Jaquess, M. J. Reinhard, M. E. Costanzo, J. A. Reggia, and R. J. Gentili, "A novel application of levenshtein distance for assessment of high-level motor planning underlying performance during learning of complex motor sequences," *Journal of Motor Learning and Development*, vol. 8, no. 1, pp. 67–86, 2019.

[5] G. E. Katz, D. Dullnig, G. P. Davis, R. J. Gentili, and J. A. Reggia, "Autonomous causally-driven explanation of actions," in *2017 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 772–778, IEEE, 2017.

[6] G. E. Katz and J. A. Reggia, "Using directional fibers to locate fixed points of recurrent neural networks," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 8, pp. 3636–3646, 2018.

[7] G. E. Katz, G. P. Davis, R. J. Gentili, and J. A. Reggia, "A programmable neural virtual machine based on a fast store-erase learning rule," *Neural Networks*, vol. 119, pp. 10–30, 2019.

[8] B. He and G. E. Katz, "Will poppy fall? predicting robot falls in advance based on visual input," in *2023 International Conference on Machine Learning and Applications (ICMLA)*, pp. 226–232, IEEE, 2023.

[9] G. E. Katz, Akshay, G. P. Davis, R. J. Gentili, and J. A. Reggia, "Tunable neural encoding of a symbolic robotic manipulation algorithm," *Frontiers in Neurorobotics*, p. 167, 2021.

[10] X. Chen, R. Liu, and G. E. Katz, "Explicit lipschitz value estimation enhances policy robustness against perturbation," in *2024 Inaugural Reinforcement Learning Conference (RLC)*, 2024. Under review.

[11] J. Ding, X. Chen, G. E. Katz, and Z. Gan, "Leveraging symmetries in gaits for reinforcement learning: A case study on quadrupedal gaits," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024. Under review.

[12] Akshay, X. Chen, B. He, and G. E. Katz, "Towards human-like learning dynamics in a simulated humanoid robot for improved human-machine teaming," in *International Conference on Human-Computer Interaction*, pp. 225–241, Springer, Cham, 2022.

[13] A. S. Salman, O. S. Salman, and G. E. Katz, "Extending cnn classification capabilities using a novel feature to image transformation (fit) algorithm," in *Intelligent Computing: Proceedings of the 2020 Computing Conference, Volume 2*, pp. 198–213, Springer International Publishing, 2020.

[14] N. Tahir, Y. Liu, T. Wang, G. E. Katz, and B. Chen, "An unsupervised approach to motion detection using wifi signals," in *2023 International Conference on Machine Learning and Applications (ICMLA)*, pp. 966–972, IEEE, 2023.

[15] G. Katz and J. Reggia, "Applications of directional fibers to fixed point location and non-convex optimization," in *International Conference on Scientific Computing*, pp. 140–146, CSREA Press, 2018.

[16] B. Sosis, G. Katz, and J. Reggia, "Learning in a continuous-valued attractor network," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 278–284, IEEE, 2018.

[17] S. Krishnagopal, G. Katz, M. Girvan, and J. Reggia, "Encoding of a chaotic attractor in a reservoir computer: A directional fiber investigation," in *2019 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2019.

[18] N. Tahir and G. E. Katz, "Numerical exploration of training loss level-sets in deep neural networks," in *2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2021.

[19] R. Liu, B. He, N. Tahir, and G. E. Katz, "On the feasibility of single-pass full-capacity learning in linear threshold neurons with binary input vectors," in *International Conference on Machine Learning (ICML)*, 2024. Under review.

[20] G. P. Davis, G. E. Katz, D. Soranzo, N. Allen, M. J. Reinhard, R. J. Gentili, M. E. Costanzo, and J. A. Reggia, "A neurocomputational model of posttraumatic stress disorder," in *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 107–110, IEEE, 2021.

[21] G. E. Katz, K. Gupta, and J. A. Reggia, "Reinforcement-based program induction in a neural virtual machine," in *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2020.

[22] G. P. Davis, G. E. Katz, R. J. Gentili, and J. A. Reggia, "Neurolisp: High-level symbolic programming with attractor neural networks," *Neural Networks*, vol. 146, pp. 200–219, 2022.

[23] G. P. Davis, G. E. Katz, R. J. Gentili, and J. A. Reggia, "Neuroceril: Robotic imitation learning via hierarchical cause-effect reasoning in programmable attractor neural networks," *International Journal of Social Robotics*, vol. 15, no. 8, pp. 1277–1295, 2023.

[24] J. A. Reggia, G. E. Katz, and G. P. Davis, "Modeling working memory to identify computational correlates of consciousness," *Open Philosophy*, vol. 2, no. 1, pp. 252–269, 2019.

[25] G. E. Katz and N. Tahir, "Towards automated discovery of god-like folk algorithms for rubik's cube," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, pp. 10210–10218, 2022.

[26] T. C. Hauge, G. E. Katz, G. P. Davis, D.-W. Huang, J. A. Reggia, and R. J. Gentili, "High-level motor planning assessment during performance of complex action sequences in humans and a humanoid robot," *International Journal of Social Robotics*, vol. 13, no. 5, pp. 981–998, 2021.

[27] X. Plourde and G. E. Katz, "Keylogging in a web-based code editor for fine-grained analysis and early prediction of student performance," in *Annual Meeting of the American Society for Engineering Education*, 2024. Under review.