**Artificial Intelligence for Synthetic Biology**

Aaron Adler, BBN Technologies, aaronadler@alum.mit.edu

Mohammed Eslami, Netrias, meslami@netrias.com

Fusun Yaman, BBN Technologies, fusun.yaman@raytheon.com

ABSTRACT:

The Artificial Intelligence (AI) for Synthetic Biology AAAI symposium was held at the Westin Arlington Gateway in Arlington, VA on October 18-20, 2018. The primary goal of the symposium was to connect and build mutually beneficial collaborations between the AI and the synthetic biology communities. The intersection of these two fields is rich with problems that can push AI beyond its current capabilities and can revolutionize discoveries and experimental design in synthetic biology. This is among the first of a series of meetings to bring these two communities together.

--------------------------------------------------------------------

Synthetic Biology integrates biology and engineering — mixing theoretical and experimental biology; engineering principles; and chemistry, physics, and mathematics. As the fields grows it is apparent that there are many opportunities and the need to apply AI techniques to several complex problem areas in the field. The Artificial Intelligence for Synthetic Biology symposium was a way to bring the two communities together. The symposium consisted of a mix of 13 technical talks, 3 invited talks, 3 discussion sessions, and a government panel.

Participants in the symposium had different backgrounds in the two fields. In order to set the foundation for the technical talks and discussions we had two invited talks: Prof. Eric Young (WPI) gave the introduction to Synthetic Biology, and Prof. Hector Munoz-Avila (Lehigh University) gave the introduction to AI. Prof. Ron Weiss (MIT) gave the keynote talk which covered programmable organoids and a discussion of modularity, which is a key idea in computer science, but not a natural property of living systems.

The symposium included a government panel that discussed funding opportunities at the AI and Synthetic Biology intersection focused on health and defense. Panelists included representatives of DARPA, NSF, NIH, and Edgewood Chemical Biological Center. Discussions included opportunities to apply AI to create, manipulate, or optimize genetic circuits, to apply AI to high fidelity experimental datasets, and to apply AI or machine learning to transfer between biological models and show that predictions are accurate.

Talks addressed existing synergies between the fields from academia, industry, and government that described variants of the design-build-test-learn cycle. AI techniques that were used included neural nets, deep learning (e.g., applied to cell free systems), active learning, Bayesian optimization, unsupervised learning techniques, planning, and semantic state models. These techniques were applied to a variety of different datasets such as instrument properties, simulated models, gene expression data, and sequence information. Challenges of applying such models to the domain were also presented. Some of the speakers, who are members of the synthetic biology industry, brought the unique the perspective of focusing on the practical and cost-effective applications of AI in optimizing synthesis of products through synthetic biology.

The symposium included three discussion sessions. The first centered around identifying the big, hard problems in synthetic biology. Identified challenges include data processing at scale, lack of quality data and metadata, outlier detection, and the need to store negative results. Other challenges centered around knowledge gaps in mapping DNA to its function, in transferring results between model systems, and in predicting biology. A lack of trust in machine learning, the need for explainability of computer suggestions, the need for knowledge in multiple fields, and the need for controlled and repeatable experiments were also discussed. The second session included highlights of AI expertise of attendees and how the problems from the first discussion might be addressed by AI techniques. The discussion centered around the need for high-quality data. Data collection, repositories, standards, and incentives were discussed, along with suggestions for various test cases. The final discussion centered on ethics surrounding AI and synthetic biology. Discussion topics included dual use concerns, gender balance of the fields, boundaries or parameters for research, and ensuring that there is sufficient upstream public engagement.

The symposium concluded with a discussion of next steps, target publications, and future meeting venues. Aaron Adler (BBN Technologies), Mohammed Eslami (Netrias, LLC), Jesse Tordoff (MIT), and Fusun Yaman (BBN Technologies) served as co-chairs of the symposium. Some papers and talks are available on the symposium website: https://www.synbiotools.com/ai-for-synbio-fss-2018/ .

Aaron Adler is a Senior Scientist at BBN Technologies in Columbia, MD.

Mohammed Eslami is the Chief Data Scientist at Netrias, LLC, in Arlington, VA.

Fusun Yaman is a Senior Scientist at BBN Technologies in Cambridge, MA.