Undergraduate Thesis Prospectus

Review of Static Analysis Methods for Probabilistic Programs

(technical research project in Computer Science)

Overreliance on Algorithms: Competition over Medical AI in Clinical Care

(sociotechnical research project)

by

Eric Thomas

November 2, 2020

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.
Eric Thomas

Technical advisor: Aaron S. Bloomfield, Department of Computer Science

STS advisor: Peter Norton, Department of Engineering and Society

General Research Problem

How can algorithmic bias be reduced?

As institutions augment professionals with computers in decision making, they introduce complex problems of responsibility (Schildt, 2016). In algorithms that guide decisions that matter to humans, bias is inevitable (Rainie & Anderson, 2017). Algorithms can improve expert decision making, but undetected biases in algorithms can deepen social divisions and inequities (Rainie & Anderson, 2017).

Review of Static Analysis Methods for Probabilistic Programs

How can probabilistic programs be analyzed without executing them?

This solo CS-capstone project will be advised by Aaron S. Bloomfield of the CS department, which will not begin until next spring.

Static analysis attempts to perform as much as possible at compile-time to gain some benefit every time the program is run (Bernstein, 2019). While dynamic analysis (i.e., during runtime) is effective at finding errors, static analysis excels at exploring all possible execution paths and variable values (Intel, 2020). Static analysis allows for compiler optimizations that shave seconds off the execution of programs that will be run a million times (Bernstein, 2019). To verify the lack of malfunctions, NASA engineers used static analysis to prove that the Mars rover will never divide by zero (Bernstein, 2019). By identifying mistakes in a program, a code editor can potentially save hours of development time, all thanks to static analysis (Bernstein, 2019). Since probabilistic programs are dependent on uncertain data, the behavior of such software can become too complex and unpredictable for traditional static analysis, requiring new techniques.

By analyzing recently published peer-reviewed articles and reimplementing the authors' work, this project will aim to find new questions, flaws, insights, and directions for future research.

Bernstein (2019) states that static analysis for probabilistic programs is currently done through optimization, verification, and usability. Optimization includes avoiding parts of the program that do not contribute to the final outcome, computing as much as possible at compile-time rather than runtime, and transforming programs to be more efficient for a certain posterior-inference method (Bernstein, 2019). Verification includes sampling paths in the program to create bounds on the probability of a predicate (i.e., a function that returns true or false) and creating an abstract program in the predicate domain (i.e., a collection of all possible values that the predicate may take) to perform posterior inference on (Bernstein, 2019). Usability includes presenting easier and more powerful interfaces for probabilistic programming (Bernstein, 2019). Since static analysis of probabilistic programs is relatively new, Bernstein suggests there is more work to be done on optimization, verification, and usability, such as transforming such programs to take advantage of parallel computing, simplifying the environment for abstract programs, and designing a restrictive mode for probabilistic programming languages to reduce the likelihood of mistakes, respectively.

Since this project focuses on reimplementation of existing research, the articles' methodologies will be tested with software to identify potential improvements. If successful, new findings that the authors may have overlooked will be identified, establishing a clearer path for future work to build from.

Overreliance on Algorithms: Competition over Medical AI in Clinical Care

How do physicians, hospitals, insurers, patient advocacies, and med-tech vendors compete to draw the line between legitimate and excessive reliance on medical AI in clinical care?

Although medical AI has the potential to reduce medical errors and group medical knowledge into a more accessible source for clinicians, AI can also magnify human biases and discrimination, resulting in differences in medical care that vary depending on the patient's gender, race, and sexual orientation (Angehrn et al., 2020; Asan et al., 2020; Ward, 2019). Because AI could harm many patients, the healthcare sector must address patient advocacies' concerns and define responsible AI use.

Asan et al. (2020) found that an implementation of machine learning for detecting skin cancer was trained on less than 5 percent of data from dark-skinned patients, which could lead to racial bias. Angehrn et al. (2020) concluded that the primary obstacles to medical AI are external validation, data exchange and privacy, and implementation logistics. By reviewing survey results, Esmaeilzadeh (2020) showed that technological, ethical, and regulatory concerns affect the perceived risks of AI use. Asan et al. (2020) suggest that trust in AI can be improved by increasing transparency, ensuring robustness, and encouraging fairness. They argue that the FDA should rigorously benchmark and test the AI to ensure meaningful outcomes and reliable surrogate endpoints. Notably, clinicians are still held responsible if they follow an AI recommendation that deviates from the standard care procedure, which can impact the clinicians' trust in AI. Since quantitative measures for the ideal level of trust between clinicians or patients and AI remain unknown, regulation on AI, such as from the FDA, continues to be lacking.

The insurer Optum used algorithms to find patients who need more resources, but they determined, falsely, that black patients were less sick (Paul, 2019). In its defense, Optum contended that "these tools should be continually reviewed and refined" and that the algorithms “should never be viewed as a substitute for a doctor’s expertise and knowledge” (Paul, 2019). Optum says “we are only scratching the surface of how they [predictive algorithms] will help improve healthcare,” implying that they believe AI use should be expanded and improved (Morse, 2019).

Alliance for AI collaborates with industry and regulators to “work to develop standards that will ensure trustworthiness and transparency in decisions supported by Intelligent Agents [AI]” (Alliance, 2019). Consistent with the recommendations of Asan et al. (2020), the Alliance offer guidelines to improve trustworthiness in AI, such as analyzing data before building models, identifying outliers in the data, validating the quality of public datasets, and communicating AI standards to society (Allgood et al., 2019). They suggest that we should “evaluate key factors that may affect our judgment of trustworthiness,” such as industry standards, accuracy, security, data privacy, ethical standards, and standardization of AI solutions (Allgood et al., 2019). While the Alliance advocates for standards and guidelines for AI, they concede that it “can enable scientists, clinicians, and medical professionals . . . to make better decisions” (Alliance, 2019).

The Society for Women’s Health Research, a nonprofit, contends that healthcare professionals must be “considering ethics at every stage of data analysis and automated decision-making” (Erickson, 2020). Some in the tech sector concur, “calling for greater awareness of ethical issues, drafting ethical codes for data science, and researching ways to identify and mitigate bias in AI.” Furthermore, they point out that bias in AI, and the medical field in general, stems partly from how “health research has tended to study white men, and until 1993, women of childbearing age were actively excluded from most clinical trials.” Since only 12 percent of machine-learning researchers are women, they assert “lack of diversity in the field of AI likely contributes to this problem.” Nonetheless, they are “optimistic about the field’s ability to overcome the perils of AI and harness its promise for women’s health” (Erickson, 2020).

The physician Verghese (2018) remarks that today “the patient in the hospital bed is . . . a place holder for the real patient who is not in the bed but in the computer.” While he states “starting with good data is critical for medical applications of AI,” medical data tends to be messy. As a result, he says “what AI will provide is at best a recommendation that a physician using clinical judgment must decide how to apply.” After noting that AI must be thoroughly vetted, he argues how “technology that is not subject to such scrutiny doesn’t deserve our trust, nor should we ever allow it to be deeply integrated into our work.”

IBM’s profile as a med-tech vendor rose when it supplied its AI, Watson, to clients in healthcare. Yet IBM admits that “we often inadvertently encode human prejudice, bias, and wrong decisions into our algorithms” (Escherich, 2019). Therefore, it suggests that “we need to start a discussion of how to best utilize the algorithms, how to understand if they contain bias, whether we are using them correctly, and whether we are setting boundaries in the right places” (Escherich, 2019). It proposes what it calls “precision regulation” for companies, which includes five policies: designate a lead AI ethics official, define different rules for different risks, make the purpose of AI clear to consumers, document information for consumers, and test AI for bias (IBM, 2020). Despite saying “the best way to promote transparency is through disclosure,” its proposal “does not entail companies revealing source code or other forms of trade secrets or IP” (IBM, 2020).

References

Allgood, B., Rodriguez, O., Bedorf, J. et al. (2019, September). *Artificial Intelligence in Healthcare: A Technical Introduction*. Alliance for Artificial Intelligence in Healthcare. https://www.theaaih.org/publications/a-primer-on-ai-in-healthcare

Alliance for Artificial Intelligence in Healthcare. (2019). *Executive Summary: AI in Healthcare*. https://www.theaaih.org/publications/blog-post-title-two-xx527-fhceb-kelx3-glmef

Angehrn, Z., Haldna, L., Zandvliet, A. S. et al. (2020). Artificial Intelligence and Machine Learning Applied at the Point of Care. *Frontiers in Pharmacology*, *11*, Article 759. doi:10.3389/fphar.2020.00759

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. *Journal of Medical Internet Research*, *22*(6), Article e15154. doi:10.2196/15154

Bernstein, R. (2019, September 12). *Static Analysis for Probabilistic Programs*. Columbia University. https://arxiv.org/pdf/1909.05076

Erickson, L. (2020, February 18). *The Promise and Peril of AI in Women's Health*. Society for Women's Health Research. https://swhr.org/the-promise-and-peril-of-ai-in-womens-health/

Escherich, K. (2019, October 7). *Why Do We Need to Talk About Ethics and Bias in AI?* IBM. https://www.ibm.com/blogs/nordic-msp/ethics-and-bias-in-ai/

Esmaeilzadeh, P. (2020). Use of AI-Based Tools for Healthcare Purposes: A Survey Study From Consumers' Perspectives. *BMC Medical Informatics and Decision Making*, *20*(1), Article 1191. doi:10.1186/s12911-020-01191-1

IBM. (2020, January 21). *Precision Regulation for Artificial Intelligence*. https://www.ibm.com/blogs/policy/ai-precision-regulation/

Intel. (2020). *Intel Inspector User Guide*. https://software.intel.com/content/www/us/en/develop/documentation/inspector-user-guide-windows/top/getting-started/dynamic-analysis-vs-static-analysis.html

Morse, S. (2019, October 25). *Study Finds Racial Bias in Optum Algorithm*. Healthcare Finance. https://www.healthcarefinancenews.com/news/study-finds-racial-bias-optum-algorithm

Paul, K. (2019, October 25). *Healthcare Algorithm Used Across America Has Dramatic Racial Biases*. The Guardian. https://www.theguardian.com/society/2019/oct/25/healthcare-algorithm-racial-biases-optum

Rainie, L., & Anderson, J. (2017, February 8). *Code-Dependent: Pros and Cons of the Algorithm Age*. Pew Research Center. https://www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/

Schildt, H. (2016). Big Data and Organizational Design – the Brave New World of Algorithmic Management and Computer Augmented Transparency. *Innovation: Organization & Management*, *19*(1), 23-30. doi:10.1080/14479338.2016.1252043

Verghese, A. (2018, May 16). How Tech Can Turn Doctors Into Clerical Workers. *The New York Times*. https://www.nytimes.com/interactive/2018/05/16/magazine/health-issue-what-we-lose-with-data-driven-medicine.html

Ward, L. (2019, October 14). The Ethical Dilemmas AI Poses for Health Care. *The Wall Street Journal*. Factiva.