AI Matters

Annotated Table of Contents

constrained mobile robots

Links

SIGAI website: http://sigai.acm.org/
Newsletter: http://sigai.acm.org/aimatters/
Blog: http://sigai.acm.org/ai-matters/
Twitter: http://twitter.com/acm_sigai/
Edition DOI: 10.1145/3203247

Join SIGAI

Students $11, others $25
For details, see http://sigai.acm.org/
Benefits: regular, student

Also consider joining ACM.

Our mailing list is open to all.

Notice to Contributing Authors to SIG Newsletters

Submit to AI Matters!

We're accepting articles and announcements now for the next issue. Details on the submission process are available at http://sigai.acm.org/aimatters.

AI Matters Editorial Board

Contact us: aimatters@sigai.acm.org

Contents Legend


Book Announcement


Ph.D. Dissertation Briefing


AI Education


Event Report


Hot Topics


Humor


AI Impact


AI News


Opinion


Paper Précis


Spotlight


Video or Image

Details at http://sigai.acm.org/aimatters

Welcome to AI Matters 4(2)

Amy McGovern, Editor (University of Oklahoma; aimatters@sigai.acm.org)
DOI: 10.1145/3236644.3236645

Issue overview

Welcome to the second issue of the fourth volume of the AI Matters Newsletter. This issue focuses on AI and Ethics, which is very timely given how often these subjects are appearing in the news lately. What really is fairness? How does an algorithm make a decision that can affect human's lives forever? Can it make a decision about whether or not to kill a person, in the case of autonomous weapons or in the case of autonomous cars? These and other other issues are discussed in the Larry Medsker's Policy column, Abhinav Maurya's report on the IEEE Big Data panel discussion about bias and transparency, Toby Walsh's opinion piece of AI as related to robber barons, and Bianca Helena Ximenes' contributed paper on autonomous car decision making.

Remember our request for new contributing editors last issue? We welcome several new contributing editors with this issue! Abhinav Maurya contributed the event report on the IEEE Big Data panel, Cameron and Tracey Hughes contributed an article on AI Cosmology, and Bianca Helena Ximenes contributed a paper on autonomous car/trolly decision making. You will be seeing more from them soon, I'm sure! Also, Sriraam Natarajan helped with overall editing for this issue.

In our regular columns, Todd Neller, our Education columnist, provides a great analysis of winning approaches to a recent Kaggle contest on predicting click-through rates. Michael Rovatsos contributed two columns, one on upcoming events of interest to the SIGAI community and one summarizing recent AI events. Be sure to check both of those out! Finally, we have a dissertation summary from Tauhidul Alam. This is a good time to remind the SIGAI community that we always welcome dissertation summaries. We also welcome a summary of the challenges of outreach from ACM SIGAI's Jinhong K. Guo and Rosemary Paradis.

Postdoc Networking Tour in Germany

We would like to point our readers to an initiative of the German Academic Exchange Service (DAAD, www.daad.de). This year, they offer a "Postdoctoral Researchers? Networking Tour" for postdocs working on Artificial Intelligence topics who are interested in (academic and non-academic) job opportunities in Germany. They offer

- a practice-oriented and diverse program to suit the requirements of the participants;
- coverage of program-related costs in Germany (accommodation, domestic travel, and most meals);
- a lump sum travel allowance if such costs are not covered by a third party, which will also be provided if the stay is extended by up to five working days in order to allow for additional professional networking.

The dates of the tour are September 23-29, 2018. The application deadline is July 15, 2018. Detailed information is available at: https://www.daad.de/veranstaltungen/networking-tour-2018/de/64036-postdoctoral-researchers-networking-tour-01/.

Submit to AI Matters!

Thanks for reading! Don't forget to send your ideas and future submissions to *AI Matters*! We're accepting articles and announcements now for the next issue. Details on the submission process are available at http://sigai.acm.org/aimatters.

Amy McGovern is chief Editor of AI Matters. She is a Professor of computer science at the University of Oklahoma and an adjunct professor of meteorology. She directs the Interaction, Discovery, Exploration and Adaptation (IDEA) lab. Her research focuses on machine learning and data mining with applications to high-impact weather.

AI Education Matters: Lessons from a Kaggle Click-Through Rate Prediction Competition

Todd W. Neller (Gettysburg College; tneller@gettysburg.edu)

DOI: 10.1145/3236644.3236646

Introduction

In this column, we will look at a particular Kaggle.com click-through rate (CTR) prediction competition, observe what the winning entries teach about this part of the machine learning landscape, and then discuss the valuable opportunities and resources this commends to AI educators and their students.

Kaggle's Criteo Display Advertising Challenge

Kaggle[1] is a data science/statistics/machine learning website that offers an excellent platform for modeling and prediction competitions. Data for training and analysis is often provided by companies, and top performers in competitions are encouraged and often required to supply and document their winning entries, offering valuable snapshots to current best practices in varied machine learning and data mining tasks.

Four years ago, Criteo Labs ran a Kaggle competition concerning CTR prediction called the "Criteo Display Advertising Challenge"[2]. The February 10, 2014 Criteo dataset was no longer available via the Kaggle competition site, but is still currently available from Criteo Labs[3]. The dataset is described on the Kaggle competition site as follows:

File descriptions

`train.csv` The training set consists of a portion of Criteo's traffic over a period of 7 days. Each row corresponds to a display ad served by Criteo. Positive (clicked) and negatives (non-clicked) examples have both been subsampled at different rates in order to reduce the dataset size. The examples are chronologically ordered.

`test.csv` The test set is computed in the same way as the training set but for events on the day following the training period.

Data fields

Label Target variable that indicates if an ad was clicked (1) or not (0).

I1-I13 A total of 13 columns of integer features (mostly count features).

C1-C26 A total of 26 columns of categorical features. The values of these features have been hashed onto 32 bits for anonymization purposes. The semantic of the features is undisclosed.

The training set consists of 45,840,617 examples, so competitors had to consider the size of the data when approaching the problem. The number of unique categorical feature values, for example, meant that a normal one-hot encoding of categorical features was computationally infeasible. Many numeric feature distributions were significantly skewed, so discretization via equal-width binning was inadvisable.

Also significant was the number of missing values in the dataset. Many machine learning (ML) and statistical learning texts have little or no coverage of the handling of missing values, and my own ML game research applications often involve complete information, so this wrinkle in both numeric and categorical data provides opportunities for learning beyond familiar, clean datasets.

Lessons from the Winners

Winners of this and 3 other recent CTR prediction competitions most often used two types of algorithms: gradient-boosted trees (GBTs, e.g. XGBoost Chen & Guestrin (2016)[4]), and field-aware factorization machines (FFMs, e.g.

[1] https://www.kaggle.com

[2] https://www.kaggle.com/c/criteo-display-ad-challenge

[3] https://s3-eu-west-1.amazonaws.com/criteo-labs/dac.tar.gz

[4] https://xgboost.readthedocs.io

libffm[5] Juan et al. (2016)). Even the winning team of Criteo's challenge made use of gradient-boosted decision trees to generate features for their FFMs[6].

Decision trees, a.k.a. classification and regression trees (CARTs), can handle missing values with ease, so the shortest path for a practitioner to see success in CTR prediction or related problems would be to learn the use of XGBoost. As an AI educator, I would want my students to *understand* GBTs, so I would want to guide them through the concept dependencies leading up to the understanding of GBTs.

In a previous column (Neller (2017)), I recommended general machine learning teaching resources for introducing the general problem of supervised learning. In that context, provide a basic introduction to decision trees using one of many good references (e.g. Quinlan (1986), James et al. (2014), §8.1, Russell & Norvig (2009), §18.1-18.3, Murphy (2012), §16.1-16.2, Mitchell (1997), Ch. 3). Next, introduce the concept of boosting (e.g. James et al. (2014), §8.1, Hastie et al. (2009), Ch. 10) and then gradient boosting (e.g. Hastie et al. (2009), §10.10, Chen & Guestrin (2016)).

Given the dominance of Python in the Kaggle community[7], I would recommend pairing these readings with practical Python exercises through Kaggle machine learning tutorials[8], the well-crafted, ongoing introductory competition on survivor prediction given Titanic passenger data[9], and even working with a subset of the Criteo dataset. I would further note that Kaggle now offers Kaggle InClass[10], a free, self-service platform that allows instructors to create classroom competitions.

As I explored Kaggle's Criteo CTR prediction competition and considered how I would guide a student to an appreciation of that work, I gained a great appreciation for the many authors that provide a foundational understanding for boosting trees, the excellent Kaggle data science community and their amazing platform, and the companies that partner with Kaggle to bring interesting challenges for the great educational benefit of all. I hope this column sparks your curiosity to explore the exciting educational opportunities these abundant resources offer.

[5] https://github.com/guestwalk/libffm

[6] https://www.csie.ntu.edu.tw/~r01922136/kaggle-2014-criteo.pdf

[7] https://www.kaggle.com/surveys/2017

[8] https://www.kaggle.com/learn/machine-learning, XGBoost-specific tutorial: https://www.kaggle.com/dansbecker/learning-to-use-xgboost

[9] https://www.kaggle.com/c/titanic

[10] https://www.kaggle.com/about/inclass/overview

References

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2939672.2939785 doi: 10.1145/2939672.2939785

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference and prediction* (2nd ed.). Springer.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014). *An introduction to statistical learning: with applications in R*. Springer. (http://www-bcf.usc.edu/~gareth/ISL/)

Juan, Y., Zhuang, Y., Chin, W.-S., & Lin, C.-J. (2016). Field-aware factorization machines for CTR prediction. In *Proceedings of the 10th ACM Conference on Recommender Systems* (pp. 43–50). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2959100.2959134 doi: 10.1145/2959100.2959134

Mitchell, T. M. (1997). *Machine learning* (1st ed.). New York, NY, USA: McGraw-Hill, Inc.

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. The MIT Press.

Neller, T. W. (2017, July). AI education: Machine learning resources. *AI Matters*, *3*(2), 14–15. Retrieved from http://doi.acm.org/10.1145/3098888.3098893 doi: 10.1145/3098888.3098893

Quinlan, J. R. (1986, March). Induction of decision trees. *Mach. Learn.*, *1*(1), 81–106.

Retrieved from http://dx.doi.org/10.1023/A:1022643204877 doi: 10.1023/A:1022643204877

Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Upper Saddle River, NJ, USA: Prentice Hall.

Todd W. Neller is a Professor of Computer Science at Gettysburg College. A game enthusiast, Neller researches game AI techniques and their uses in undergraduate education.

Events

Michael Rovatsos (University of Edinburgh; mrovatso@inf.ed.ac.uk)
DOI: 10.1145/3236644.3236647

This section features information about upcoming events relevant to the readers of AI Matters, including those supported by SIGAI. We would love to hear from you if you are are organizing an event and would be interested in cooperating with SIGAI, or if you have announcements relevant to SIGAI. For more information about conference support visit sigai.acm.org/activities/requesting_sponsorship.html.

15th International Conference on Informatics in Control, Automation, and Robotics (ICINCO 2018)

Porto, Portugal, July 29-31, 2018
http://www.icinco.org/
The purpose of the 15th International Conference on Informatics in Control, Automation and Robotics (ICINCO) is to bring together researchers, engineers and practitioners interested in the application of informatics to Control, Automation and Robotics. Four simultaneous tracks will be held, covering Intelligent Control Systems, Optimization, Robotics, Automation, Signal Processing, Sensors, Systems Modelling and Control, and Industrial Informatics. Informatics applications are pervasive in many areas of Control, Automation and Robotics; This conference intends to emphasize this connection.
Submission deadline: June 5, 2018

Foundations of Digital Games 2018 (FDG 2018)

Malmö, Sweden, August 7-10, 2018
fdg2018.org
The International Conference on the Foundations of Digital Games (FDG) is a major international event. It seeks to promote the exchange of information concerning the foundations of digital games, technology used to develop digital games, and the study of digital games and their design, broadly construed. The goal of the conference is the advancement of the study of digital games, including but not limited to new game technologies, critical analysis, innovative designs, theories on play, empirical studies, and data analysis. FDG 2018 will include presentations of peer-reviewed papers (with rebuttal process), invited talks by high-profile industry and academic leaders, panels, workshops, and posters. The conference will also host a game competition, tech demo session, and a doctoral consortium. This years FDG conference will nominate two papers with honorable mention and one best paper from each track. FDG 2018 is organized in-cooperation with ACM SIGAI, SIGCHI, and SIGGRAPH.
Submission deadline: March 22, 2018

33rd IEEE/ACM International Conference on Automated Software Engineering (ASE 2018)

Montpellier, France, September 3-7, 2018
www.ase2018.com
The IEEE/ACM Automated Software Engineering (ASE) Conference series is the premier research forum for automated software engineering. Each year, it brings together researchers and practitioners from academia and industry to discuss foundations, techniques, and tools for automating the analysis, design, implementation, testing, and maintenance of large software systems.
Submission deadline: April 26, 2018

2nd ACM Computer Science in Cars Symposium (CSCS 2018)

Munich, Germany, September 13-14, 2018
https://cscs.mpi-inf.mpg.de
Industry as well as academia have made great advances working towards an overall vision of fully autonomous driving. Despite the success stories, great challenges still lie ahead of us to make this grand vision come true. On the one hand, future systems have to be yet more capable to perceive, reason and act in complex real world scenarios. On the other hand, these future systems have to comply with our expectations for robustness, security and safety.

ACM, as the worlds largest computing society, addresses these challenges with the ACM Computer Science in Cars Symposium. This conference provides a platform for industry and academia to exchange ideas and meet these future challenges jointly. The focus of the 2018 conference lies on AI and Security for Autonomous Vehicles. Contributions centered on these topics are invited.
Submission deadline: May 28, 2018

10th International Joint Conference on Computational Intelligence (IJCCI 2018)

Seville, Spain, September 18-20, 2018
http://www.ijcci.org
The purpose of the International Joint Conference on Computational Intelligence – IJCCI – is to bring together researchers, engineers and practitioners interested on the field of Computational Intelligence both from theoretical and application perspectives. Four simultaneous tracks will be held covering different aspects of Computational Intelligence, including evolutionary computation, fuzzy computation, neural computation and cognitive and hybrid systems. The connection of these areas in all their wide range of approaches and applications forms the International Joint Conference on Computational Intelligence.
Submission deadline: June 13, 2018

5th international Workshop on Sensor-based Activity Recognition and Interaction (iWOAR 2018)

Berlin, Germany, September 20-21, 2018
https://www.iwoar.org
This conference-like workshop is initiated and organized by the Fraunhofer IGD and the University of Rostock. It offers scientists, interested parties, and users in the area of sensor-based activity recognition and interaction the possibility to an exchange of experiences and a presentation of best-practice examples, as well as technical and scientific results. The workshop focuses on technologies for human activity recognition and interaction via inertial sensors (accelerometers, gyroscopes etc.) and their scientific applications.
Submission deadline: June 15, 2018

Michael Rovatsos is the Conference Coordination Officer for ACM SIGAI, and a faculty member of the School of Informatics at the University of Edinburgh, UK. His research in in multiagent systems and human-friendly AI. Contact him at mrovatso@inf.ed.ac.uk.

Conference Reports

Michael Rovatsos (University of Edinburgh; mrovatso@inf.ed.ac.uk)
DOI: 10.1145/3236644.3236648

This section features brief reports from recent events sponsored or run in cooperation with ACM SIGAI.

23rd ACM International Conference on Intelligent User Interfaces (IUI 2018)

Tokyo, Japan, March 7-11, 2018
http://iui.acm.org/2018/

This was the 23rd IUI conference, continuing the tradition of being the main forum for reporting outstanding research at the intersection of HCI and AI. The work that appears at IUI bridges these two fields and delves also into related fields, such as psychology, cognitive science, computer graphics, the arts, and others. The program of IUI 2018 reflected the growth of the Intelligent User Interfaces research community. The calls for contributions attracted 297 full and short paper submissions from all over the world (a record for IUI conferences), 127 submissions of posters and demos, and 22 submissions to the student consortium. The conference committee accepted 68 papers (43 long papers and 25 short papers), covering a diverse range of topics. The program also included 35 posters, 30 demos, and 11 student consortium papers. In addition, IUI 2018 featured 7 workshops on topics related to Intelligent User Interfaces. One of the main features of the conference were the 3 keynote talks. James Landay from Stanford University opened the conference with a keynote entitled “From On Body to Out of Body User Experience.” Following this, Masataka Goto from the National Institute of Advanced Industrial Science and Technology (AIST) presented his talk “Intelligent Music Interfaces.” Finally, Jennifer Golbeck from the University of Maryland presented her keynote“Surveillance or Support: When Personalization Turns Creepy.” IUI 2018 also feature the second edition of the Impact Award, celebrating an impactful paper presented at the past editions of IUI. A novel aspect of IUI 2018 was its co-location with IPSJ Interaction 2018, the leading domestic HCI conference in Japan. The two conferences were held back-to-back in the Hitotsubashi Hall, allowing the participants of one conference to also take part in the other. The two conferences had a shared day feature a keynote talk and a shared interactive poster/demo session. Overall, this was the largest and one of the most successful IUI conferences, attracting close to 400 participants..

11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018)

Funchal, Portugal, January 19-21, 2018
http://www.biostec.org/?y=2018

BIOSTEC received 340 paper submissions from 49 countries and was attended by 265 participants. The conference programme included paper presentations (acceptance rate 19.71%) as well as four invited talks by Anatole Lcuyer (Inria Rennes/IRISA, Hybrid Research Team, France), Corina Sas (Lancaster University, United Kingdom), Dinesh Kumar (RMIT University, Australia), and Maximiliano Romero (Università Iuav di Venezia, Italy). The BIOSTEC program also included a special session on Knowledge Acquisition and Learning in Semantic Interpretation of Medical Image Structures (BIOIMAGING) lecture by Piotr Szczepaniak, Piotr Grzelak and Arkadiusz Tomczyk, a Special Session on Assessing Human Cognitive State in Real-World Environments (BIOSIGNALS) lectured by Bethany Bracken and a Special Session on Neuro-electrostimulation in Neurorehabilitation Tasks (BIOSIGNALS) lectured by Vladimir Kublanov. Additionally, a “Best Paper Award”, a “Best Student Paper Award” and a “Best Poster Award” were conferred at the conference.

4th International Workshop on Sensor-Based Activity Recognition and Interaction (iWOAR 2017)

Rostock, Germany, September 21-22, 2017
https://iwoar.org/2017/

iWOAR is an international workshop with conference character, which takes place at the Baltic Sea in the north of Germany. The event is initiated and organized by the Mobile Multimedia Information Systems Group at the University of Rostock and the Fraunhofer IGD in Rostock. It offers scientists, interested parties, and users in the field of sensor-based activity recognition and interaction the opportunity to exchange experiences and to present best-practice examples, as well as technical and scientific results. The workshop was attended by 32 participants, who attended the presentation, poster and discussion sessions, and two keynote talks, one by Jesse Hoey from the University of Waterloo and another by Ian Craddock from the University of Bristol. In this years edition of iWOAR there were 19 submissions, out which 12 papers were accepted, resulting in an overall acceptance rate of 63%.

9th International Joint Conference on Computational Intelligence (IJCCI 2017)

Funchal, Portugal, November 1-3, 2017
http://www.ijcci.org/?y=2017
Including four simultaneous tracks, IJCCI 2017 has been held covering different aspects of Computational Intelligence, namely evolutionary computation, fuzzy computation, neural computation and cognitive and hybrid systems. The connection of these areas in all their wide range of approaches and applications forms the International Joint Conference on Computational Intelligence. As its previous editions and since 2009, IJCCIs target audience involves both academic and industrial participants. In fact, establishing an effective connection between academic and industrial fields in Computational Intelligence is one of aims of IJCCI. Four keynotes were presented on various aspects of Computation Intelligence: António Dourado (University of Coimbra) presented a lecture dealing with "EEG Analysis and Classification for Diagnosis and Prognosis of Brain Disorders". Emma Hart (Edinburgh Napier University) presented a lecture on "Lifelong Learning in Optimization Algorithms". Paulo Novais (University of Minho) gave a lecture dealing with "Ambient Intelligent Systems and Role of Non-Intrusive and Sensitive Approaches". Finally, Jonathan Garibaldi (University of Nottingham) presented a lecture on "Type-2 Fuzzy Systems for Human Decision Making".

Foundations of Digital Games 2017 (FDG17)

Cape Cod, USA, August 14-17, 2017
fdg2017.org/
FDG17 was the 12th instantiation of the Foundations of Digital Games conference, promoting the exchange of information on the foundations of digital games, technology used to develop digital games, and the study of digital games and their design. Organized by the Society for the Advancement of the Study of Digital Games (SASDG), the goal of the conference is the advancement of the study of digital games, including but not limited to new game technologies, critical analysis, innovative designs, theories on play, empirical studies, and data analysis. FDG17 focused on Celebrating the Player. FDG18 received a total of 89 regular paper submissions to its seven tracks: game analytics and visualization, game AI, game criticism and analysis, game design and development, games for a purpose, game technology and development, player experience. Of these, we accepted 36 (acceptance rate 40%). We also accepted 24 of 31 submitted posters (acceptance rate 77%). Finally, we co-hosted three workshops with a total of 28 submitted and 19 accepted papers (acceptance rate 67%), some of which have become part and parcel of the FDG community, like the International Workshop on Procedural Content Generation, now in its 8th year.

Data Institute Conference (DSCO17)

San Francisco, CA, October 15th-17th, 2017
http://www.sfdatainstitute.org/conference.html
The Data Institute at the University of San Francisco hosted its inaugural Data Institute Conference (DSCO17) in downtown San Francisco in October 2017. The conference included over 75 invited sessions from data science experts in industry and academia. The first day of the conference offered tutorial workshop sessions in deep learning, network analysis, and experimental design. The conference also held a poster session, a panel that discussed the state of the art in data science in industry, and three plenary speakers, including Michael Jordan from

UC Berkeley, Anima Anandkumar from Amazon Web Services and Caltech, as well as JP Onnela from the School of Public Health from Harvard University. DSCO17 had over 225 attendees from data science, ranging from current Ph.D students, new and expert academic researchers, as well as data scientists from technology companies, including Google, Eventbrite, Airbnb, and Microsoft. In addition to the highlights already mentioned, one event that we the organizers were particularly proud of was the mentor lunch on the second day of the conference. The purpose of this was to pair mentors - who are considered further along in their career such as being tenured professors and senior data scientists with young and underrepresented mentees who are at the beginning of their data science career. A total of twelve mentees and thirteen mentors were present, and the lunch was a big success.

14th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2017)

Madrid, Spain, July 26-28, 2017
http://www.icinco.org/?y=2017

ICINCO 2017 was held in Madrid this year and was sponsored by the Institute for Systems and Technologies of Information, Control and Communication (INSTICC), co-organized by Universidad Rey Juan Carlos as a local partner. This edition of the conference received 214 paper submissions from 46 countries, out of which 25.7% were published and presented as full papers. Four invited talks were delivered by internationally distinguished speakers, namely Andre Rosendo (ShanghaiTech University), Vitor Santos (Universidade de Aveiro, Portugal), Wolfram Burgard (University of Freiburg, Germany), and Carme Torras (CSIC-UPC, Spain). Additionally, a ”Best Paper Award” and a ”Best Student Paper Award were conferred at the conference venue.

9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2017)

Funchal, Portugal, November 1-3, 2017
http://www.ic3k.org/?y=2017

The main objective of IC3K is to provide a point of contact for scientists, engineers and practitioners interested on the areas of Knowledge Discovery, Knowledge Engineering and Knowledge Management. To enhance exchange opportunities, we have organized an European Project Space that aims at presenting case-studies and developing partnerships between conference participants around projects in IC3K topic areas, that are financially supported by the European Community. This year, IC3K received 157 paper submissions from 47 countries. To evaluate each submission, a double blind paper review was performed by the Program Committee. After a stringent selection process, 20% of the papers were published and presented as full papers. The IC3K program also included a special session on Information Sharing Environments to foster crosssectorial and cross-border collaboration between public authorities – ISE (KMIS) and two tutorials: “How to Mine Enterprise Ontologies” lectured by Linda Terlouw and Jan Dietz and “Traceability and Structuring Knowledge from Cooperative Activity” lectured by Nada Matta. Additionally, a “Best Paper Award” and a “Best Student Paper Award” were conferred at the conference.

Michael Rovatsos is the Conference Coordination Officer for ACM SIGAI, and a faculty member of the School of Informatics at the University of Edinburgh, UK. His research in in multiagent systems and human-friendly AI. Contact him at mrovatso@inf.ed.ac.uk.

IEEE Big Data 2017 Panel Discussion on Bias and Transparency

Abhinav Maurya (Carnegie Mellon University;)
DOI: 10.1145/3236644.3236649

Panelists: Cynthia Dwork (Harvard University), **John Langford** (Microsoft Research), **Jure Leskovec** (Stanford University/Pinterest), **Jeanna Matthews** (Clarkson University)
Moderator: Ricardo Baeza-Yates (NTENT)
Scribe: Abhinav Maurya (Carnegie Mellon University)

In January 2017, the ACM US Public Policy Council released a report on algorithmic transparency and accountability (ACM US Public Policy Council, 2017) which outlined several characteristics for algorithms to be considered transparent and accountable:

- Awareness
- Access and redress
- Accountability
- Explanation
- Data Provenance
- Auditability
- Validation and Testing

A panel discussion on *Big Data Bias and Transparency* was organized at the IEEE International Conference on Big Data held in December 2017 to discuss opportunities and challenges faced by the data science community in their effort to incorporate the tenets of fairness, accountability, and transparency in their data-driven analyses and products. The panel consisted of Cynthia Dwork from Harvard University, John Langford from Microsoft Research, Jure Leskovec from Stanford University/Pinterest, Jeanna Matthews from Clarkson University, and was moderated by Ricardo Baeza-Yates from NTENT. This article provides an account of the panel discussion in the hope that it will be of interest to readers of AI Matters.

Ricardo Baeza-Yates: Perhaps the ideal form of interpretability is to have algorithms explain their decisions. Is it possible to build algorithms that can explain their decisions to convince us of their correctness?

Cynthia Dwork: I think it's difficult to pin down what a convincing explanation of a decision might be to a human. "Why was I turned down for the loan?" I have no idea how to answer that. There is a classifier, we feed your data in, this was the outcome, you were approved or turned down. A different question that I might be able to make sense of mathematically is "What is it that I can change at a reasonable cost that would lead to a different decision?" That is a question that makes some sense to me. But why was I classified this way, I can't really make sense out of it.

Jeanna Matthews: When we are talking about really important decisions like whether to send someone to jail or not, explanation might be even more important than an incredibly accurate learning algorithm. The ability to export human-readable, understandable version of important decisions makes them regulatable.

Jure Leskovec: As a community, we like taking datasets and training algorithms on them and competing on who gets the biggest AUC or F1 score. As we start thinking about really applying these methods to problems that have more consequence than whether you will click a given ad or not or maybe you will watch a given movie or not and maybe that ruins your Friday night, but that is the most serious consequence it has. When you start thinking about these more important societal applications, then the question becomes how humans and algorithms work together, and what kind of algorithms work with humans in a given way. So I think it's a much broader question than about just machine learning algorithms or systems.

John Langford: If you want your machine learning systems to be debuggable, you need to think about your model in the context of the data source. If you keep the model separate from the data source, that's a bug waiting to

happen. You need the data source attached to the model so that you can track back and discover why the model is beginning to behave in a certain way. Machine learning systems are more than just algorithms, they keep track of where the data is coming from and how it is used in training the model. For explanation, the same thing comes up. If you are trying to figure out how to create a better model, you need to explain its decisions and mistakes. Within every machine learning algorithm, there is always a bug. It is never the case that you have an optimal machine learning algorithm. There's always something you can do to improve it. Figuring why a machine learning algorithm is failing on a certain task is a great way to figure out how to improve it. So if you have ways of explaining why the model is behaving in a certain way or if you are operating in spaces where it is really obvious what the model is doing, these are mechanisms to figure out improvements to the model. Improvements to the ImageNet systems have been driven by figuring out what the bugs were in previous year's system and how to improve it. For auditability, trying to debug a non-deterministic system is really difficult. For accountability, we have a paper at the FAT-ML workshop showing that any cost-sensitive learning algorithm can be turned into a fairness-satisfying learning algorithm. Hence, if we know that there is a bias issue, we can modify our learning algorithms to address this issue systematically, at least in a much more systematic fashion than we do now.

Cynthia Dwork: (To Langford) What's your definition of fairness? Are they group definitions of fairness, or statistical parity, or something else...?

John Langford: There are several definitions. For every single definition, we can create a fairness-aware learning algorithm.

Cynthia Dwork: There are many fairness conditions that are mutually inconsistent. What happens in that case?

John Langford: What the definition of fairness should be is something that people need to figure out. But if you write down a definition of fairness and you want to have it forced, we can turn almost all definitions into a reduction which will transform a classifier into a fair learning algorithm.

Cynthia Dwork: So again I don't see how that can happen when these definitions are mutually exclusive.

John Langford: So you have to choose one. Given that you choose one, I will create a learning algorithm to give you what you want. If you don't choose one, then I can't do it.

Cynthia Dwork: And can you do it for individual fairness?

John Langford: What is the definition of individual fairness?

Cynthia Dwork: That similar people should be treated similarly. So you have some kind of metric for a given classification task which tells you how similar or dissimilar a given pair of people is. For this particular classification task, can you ensure that there is some relationship between the training distance and the probability distributions on their outcomes?

John Langford: Is this similar to equalized odds?

Cynthia Dwork: No, equalized odds is a group definition which says that this group as a whole should have similar outcome probabilities compared to other groups. But that's a group definition of fairness.

John Langford: I need to know the exact definition before I can give you an exact answer.

Ricardo Baeza-Yates: I think it might be useful to clarify what fairness is. Because politicians creating laws are unsure about it. One way to think about it is that when politicians create laws, they don't worry about the details of implementing the law. Some formalization of fairness can help politicians create rules and guidelines for programmers and businesses. My next question is on accountability of algorithms. Who is accountable for the transparency and fairness of learning algorithms? Is it the person providing the data, is it the one that programs the algorithm, is it the corporation which deploys the algorithm...? There are many implications for the future that will change the field. John, we start with you.

John Langford: The question of ethics in algorithms is related to fairness. We know that given a definition of fairness, you can train

any classification algorithm to comply with that definition, and tradeoff between accuracy and fairness. The problem is that many people don't have an actual definition of fairness or what is ethical. The second is that the algorithms have to be aware of protected attributes in order to be able to incorporate them for achieving fairness.

Ricardo Baeza-Yates: Jure, can you talk about your research with judges in this regard?

Jure Leskovec: Sure! I think these issues are really interesting and important. In our group, we have been working with the Chicago crime lab and with an economist here at Harvard. And the question we have been looking at is whether we can help criminal court judges make better bail decisions. The question is after a person is arrested, where will the person wait for trail. The person can await trial in jail, or they can be free. If free, they can misbehave - they can commit a violent crime, or they can commit a non-violent crime, or they can simply fail to appear at trial. So we were asking how can machine learning help judges make better bail decisions. It was interesting how many technical and algorithmic issues came out when we started working on this problem. Ricardo was saying that the law is very clear. And the law is that the judge should ignore the severity of the crime when making bail decisions. The judge should try to assess the probability of recidivism. As we were doing this research, one thing for example that we noticed was that machine learning algorithms could reduce the level of crime by around 40% if you keep the prison population the same. Another way to say this is if you keep the current level of crime, you could release 72-73% of the people awaiting bail. But the data collection process is itself biased. We only see the outcome of the people that were actually released, we don't know the outcomes for people who were kept in jail. If you assume that machine and human have access to the same information, there are statistical ways of imputation to get around this. But humans see much more than machines. To give an example of how bad this can be, consider that a judge learns from years of experience on sentencing young people that if their family shows up at the bail hearing, it is ok to release on bail. If the family doesn't show up, it means the person might commit another crime. Assume that we didn't go and encode this feature into our algorithm. Then, based on your release data, you will learn than young people who are released commit no crime. And any fancy or "fair" algorithm with the most proper cross-validation will tell you that young people commit no crime. And then you go and deploy this in the real world and your crime rates will go up. You suddenly see cases of young people committing crimes. What is interesting about this research is that even though the law is very clear, we developed a way to diagnose how humans make decisions and identified groups of defendants where judges make consistent decisions and on certain others they do not. For example, on single people who move around a lot and on families without kids, judges make very accurate decisions, but on people with kids, their decisions are much less accurate. One way to explain it is that this is because judges are making mistakes. Another way is to acknowledge that the objective/cost function of the judge and the algorithm are very different. If you put a single person in jail, you restrict their movement and ability to commit crime but there is little cost to the rest of the society. But if you put a person with kids in jail, there is a huge consequence on society, for the families and so on. And your decisions will now affect their future behavior. To make my long story short, very interesting things start happening if you take a real example and ask how can we build algorithms to help society and interesting aspects begin to emerge that one wouldn't even think about. And it's a very interesting area to work on.

Jeanna Matthews: So the question about ethics and societal values, those can change a lot with countries and even with different times in history. For example, in this country, we would rather let guilty people go free than put innocent people in jail. Another example is that one is not allowed to consider race in a hiring decision. I am very concerned that we are fundamentally replacing such societal values without any discussion of it. If you replace decisions like that with a piece of software, we run into problems. I think it is being labeled this way: these are unbiased logical decisions made by computers when that's not true. They are trained on historical data in which there is actually a lot of bias. Train-

ing on historical data makes sense; the past is all we have; we don't really have a choice. But we don't always want the decisions to look like the past. And when we fail to recognize it and think of it as a limitation of the tool, that's a problem. Another problem is the fundamental nature of these black-box algorithms that are being used. There are many proxies for the sensitive attributes such as race or gender. You could say that you are not looking at the sensitive attributes of race or gender, but proxies to those are what you could be looking at if you looked at the explanation. So would it be easy to build a software system that you could claim makes completely unbiased decisions and could keep certain people out of this country? Yes, you could do that and it might even be attractive in this political climate. That is very concerning. So if we have what we believe to be fundamental societal values and we are replacing human decision-making by black-box algorithmic decisions and we fail to require explanation of the decisions, we may be using sensitive attributes directly or using certain proxies which are just as good. Or it could be just a bug in the system that could lead it to make wrong decisions. In many of these systems, there is no forcing function for debugging. If you take a proprietary software system that you use to judge recidivism and the company says our intellectual property rights are more important than a defendant's right to explanation, does that sound outrageous to you? That's exactly what's happening, Louis versus Wisconsin for example. These are the things that are happening right now! We might be fundamentally changing our societal values without discussion simply by replacing human decisions with black-box decisions and without requiring explanations. I think there's a lot we can do in the technical community if we are sufficiently humble about limitations of the things we build and sufficiently advertise the dangers of using them and ways in which they are inappropriate. We should sound the alarm that they should not be used in ways that some people might want by tucking the details under the covers, sound the alarm on things that they would like to have happen by hiding under the label of completely unbiased decisions made by computers. At least, we can audit algorithms. It's more difficult to audit humans. So we have potential to do better. But we also have the potential to do a lot worse and label it as better.

Cynthia Dwork: Actually, I want to comment on a couple of things I heard before providing my answer. Thinking about fairness and the predictors that are used in legal contexts, I am not a lawyer or legal expert but I had a conversation with a PhD in Law grad student at Harvard. She was talking to me about bail decisions in New York state where the only factor one is allowed to take into account is flight risk. Ok, that's pretty concrete. But then she pointed out that there really are multiple reasons why one might be a flight risk. One is that they really might be a flight risk i.e. they might run away and not come to trial. But another is that they simply can't afford the transportation to get back to court. And incarcerating someone just because they cannot afford to get back to court seems wildly unfair. So these things are incredibly subtle and incredibly laden with context. Another thing that comes up in recidivism prediction. Imprisoning someone isn't only a question of protecting society from recidivists. There are other reasons for jailing people including punishment. Where does that get put into the mix? Having an estimate or a way of trying to estimate somehow the likelihood that somebody is going to do something violent is clearly useful but it's definitely not going to be the whole story. So this means that in order to decide sentencing, one has to sit down and decide what's the point of sentencing, and it involves enormous amount of societal context. About explanation, one of the things I hear already on the panel is these two different notions of explainability. One is explaining a particular decision, and another is akin to what you (Langford) were saying about debuggability. I dream about being able someday to say if we use this learning algorithm and these notions of fairness which I can lay out and you can examine and decide whether you like them or not, and these software principles for building systems that are fair; then maybe we could have something that is fair. I want us to get into that realm of things. We need ways that are much more systematic and catch issues besides the ones that we are already looking for. Fairness behaves oddly under composition. It does not behave like composition in cryptography or privacy-preserving data analysis. You can take two

things that are fair and you can find scenarios in which they are competing with each other, and the outcome of the system as a whole is not fair. So when we come back to me again, I will tell you a story about that.

Ricardo Baeza-Yates: The blackbox systems we deal with are so complex, and if we want to change how the system behaves, we need to understand the dynamics of the system. Also, there is the feedback loop. We collect data, make decisions, which changes the data we collect, changes the system, and so on. And the sideffects of these complex decisions cannot always be anticipated. For example, sending someone to jail might be the best way to turn someone into a criminal. You have the best training school and the best networking. So you are also changing probabilities of committing crimes in the future.

Jure Leskovec: I think this notion of exploration... If you think of bail or something else like medical procedures and so on, you cannot go ahead and collect random data. So if I want to build a skin cancer prediction system, the only way for me to collect data is for me to get a scalpel and start collecting samples, which would be amazingly non-ethical to do. I cannot come and start fooling around...

John Langford: I disagree. They do clinical trials all the time.

Jure Leskovec: No. But the point is they stop the clinical trial as soon as they have the result or they determine it is unethical. Clinical trials are not there to collect data; they are there to answer a specific question. And that's a huge difference. You cannot do random exploration. You cannot say: oh! we don't know what's happening here. Let's release this person.

John Langford: Random exploration need not be uniform or uninformed exploration. Uniform exploration is never the best kind of exploration.

Jure Leskovec: Even if it's non-random, I would say there are ethical issues with doing something that may be potentially harmful with the goal of collecting data.

Cynthia Dwork: So just to clarify, you are talking about collecting data to go into your training set? (Leskovec confirms.)

John Langford: I agree that there can be ethical issues, but I don't agree that every time you do exploration in the medical field, it is unethical. And clinical trials are a good example of this.

Jure Leskovec: Again, my point is medical trials are there to test hypotheses, not to collect data. The other thing that becomes interesting is the question of features. In bail, you have protected attributes like gender, religion, race, etc. On second thoughts, I think gender you can use but you can't use race or religion. Now, I think here's a good question that I don't have an answer to. What does it even mean not to use a protected attribute when you have lots of data and lots of correlations. Also, we were talking about families before. When we were doing our analysis e.g. how would algorithmic decisions compare to that of a human judge, the algorithm would release more black people, jail more Hispanics and jail more whites as well. Then, you can ask what if we release the same proportions of the subpopulations as the judges are releasing, and we still do better than the judge. But if we step back and ask what would be the right thing to do, we honestly don't know what should be the ratios. I think that's a big challenge - how do we think about this problem. The last anecdote that I will leave you is this. We did an experiment to understand where or why humans may be making mistakes. So we trained an algorithm that was trying to imitate the judge. So the algorithm didn't care about what's the right decision; it was just trying to imitate the judge. And when we took this artificial judge and applied it and saw what is it's accuracy, how good are the decisions that it made, this artificial judge was better than the judge it was trained on. And the only way to explain this is to say that the human judge has certain signals that the machine doesn't have access to. And whenever the machine makes a mistake - not imitating the teacher - the machine in some sense is making the correct decision. The machine didn't have access to certain signals that the judge was using in decision-making. The features that we were using in this work were based on history of criminal record - what was the age or sex, did they ever fail to appear before, and so on and so forth. These were features that were administrative, impossible to manipulate and the only way to affect them was to not commit

crime or to not get arrested I guess. There is the failure to appear, but there are also violent and non-violent crimes. We managed to reduce the violent crimes quite a bit. There is a good case why algorithmic decisions could help judges. The human judge can only see so many cases, the algorithm can see millions. When we were talking to judges, they told us that they have 30 seconds to a minute to make a decision. And after they make a decision, it is nearly impossible for them to see the outcome of the decision. They told us that the only way for them to learn whether it was a correct decision was to check the local newspapers and see if the released committed any crimes or not. So it's a very, very hard problem for the judges.

Jeanna Matthews: So the best criminals may not get arrested or may be people lucky enough to not live in jurisdictions with a high rate of arrest for crimes. Now, there are two points I would like to make. One is about the accuracy of data. We all know that in this world of big data, there is a lot of messy inaccurate data. So that's another important aspect of explanation. I cannot tell you how often I have looked at summaries like mean of the data and thought that is absolutely not true. I think there was a case recently about a proprietary recidivism software where somebody was saying one of the input pieces was incorrect, and they were arguing that it needed to be repeated. So it is not just consumer data but also your go-to-court kind of data where there are inaccuracies. So that's one thing. The other thing that I want to say is what are the forcing functions for debugging. I had a chance to go visit the Legal Aid Society of New York and they were talking about forensic software that are used to perform DNA matches of their clients. There was a software package where they got access to the source code, and which is now on Github. They found some very weird examples of bugs. One was where it should have been the case that the code erred on the side of not matching people and they found bugs where that was not the case. Also, that set of companies fight tooth and nail not to have disclosure of their software in any way in court, not even in a protected way, like not even the legal counsel gets to see it. Individual defense teams have to fight to get access to analyze the software. That's kind of a crazy world to be living in. It's very difficult to get the right to do that. And when you get in and see it, you find it's not doing what it's supposed to be doing. Maybe the data is inaccurate or messy. And in this world of criminal justice, maybe someone says I am not guilty, I swear I am not. And the system says, of course you say you are not, but you are a match and are going to jail. Again, I would like to ask what's the forcing function for debugging. Some of the defendants are true when they say they didn't do it. Are we going to lump all that together? If you have ever used a random software package, you know there are tons of bugs. You know there are bugs in there, right? What if you had to live with them forever because every time someone tried to report a bug, it was just dismissed? More importantly, what is the incentive of these companies to debug, to improve or make things better. They might feel that their software is perfectly fine. There might even be some buyers who are happier if it notches up the guilty ratings. They might be perfectly happy with the system as it is. They don't need any debugging, and don't need more accuracy or testing. There are some people going to jail. Our constituents are happy with that. We are good here... Until it is you or your family or friends. And also what population demographic is it more likely to be? There are just a lot of issues there.

Cynthia Dwork: I find this absolutely fascinating. I have a question. In cases where a mistake was found, is it something that required examination of the software or is it something where you already knew the answer and you were checking what the system output would be?

Jeanna Matthews: The specific case they were talking about involved source code analysis and finding a routine that did something which the software creators swore it did not do. I get why these companies might not want to reveal their software. But one of the more dangerous cases for me would be a company that said you can look at our system, it's completely open, but the problem is in the training data. Let me be a little more organized in my thought. One, you might have a problem in the data, not in the software at all. Two, companies don't want to reveal their software, so maybe we don't have to fight that battle.

Maybe what's better is targeted testing, being able to tweak things and see how the output changes. And some of the companies do provide some of these features. You could change little things and see what the answer would be. In the judicial context, they would prefer that than letting people look into the software. But then you would have to trust them, the answers they give back, this is the state of the system. And again the question I would ask is: what is the forcing function for them to make the system better and better and truly find bugs when they are just as happy without the extra effort? We all want to think that the systems we build are perfect and good and don't have any bugs. But apart from that what is the forcing function to find bugs, especially the harder corner case bugs that escape even source code examination.

John Langford: So with respect to such software, I think there should just be a law promoting transparency and open-source software. I don't see how we can trust a black-box to handle each case correctly.

Jeanna Matthews: I agree on open source for public use software. It's just that there are intellectual property rights issues that prevent open sourcing all proprietary software...

John Langford: But there is much more to a system than the algorithm. The algorithm can be made public and examined without all the system details.

Cynthia Dwork: Do you feel the same way about medical diagnostic devices that have circuits in them?

John Langford: I might...

Cynthia Dwork: It's interesting that we don't hear so much discussion of it.

John Langford: So another good example of bad data was Senator Patrick Leahy who discovered fake comments attributed to him that were anti-net neutrality even though he is pro-net neutrality. So there's a lot of bad data there.

The panel discussion turned at this point to answering audience question before finally summarizing the key takeaways from the discussion. The takeaways from each of the panelists are recorded below.

Jeanna Matthews: I will just reiterate that I think in some ways we might be changing our societal tradeoffs without any discussion by replacing some of our current processing with black-box decisions. That is something we should educate people about and care about. The potential for mischief in black-box systems is very high. I think we want to debug our systems, but not everyone who builds these systems may want to debug them and share our goals of transparency. If we begin to accept black-box decisions as being better than human decisions, that is a very dangerous road to go down on. Even if there is a cost in terms of accuracy, if we are talking about regulatable decisions, it's important to insist on explanations because the potential for mischief and bugs is too high, and the history of that kind of stuff is not good.

John Langford: I think we need a wider debate with society. I think there are two characteristics that make our current black-box decisions prone to bias: first is that they are black-box and second is that they are currently untestable. Sometimes, black-box systems are testable and that is enough, but if it's both black-box and untestable, then it's just ridiculous.

Jure Leskovec: My view would be that computer scientists or machine learning people or data people should actually go out there and be part of the debate and do real work. We can keep talking about this in our immediate community and write our papers, but the value of this is limited. When we get out of this conference zone and work on our concrete problem on a concrete application, people will care about that. This way I think we will learn much more about problems - what is a real problem and what is a made-up problem. We can then drive the agenda going further. What we learned in our research is that it's important to go out and say how can we do this better. Expose yourself, go out of our comfortable circles, and attack problems in the real world. This way new problems will arise, and we will solve them. We have to solve them, because none else will.

Cynthia Dwork: So I think it's a really good point. We have a lot of responsibility. Policy people don't understand the issues enough. When they are educated, my experience is that they turn around and say "ok, so what

do we do now." We just can't avoid trying to come up with answers. It's not that we have to get the final answers, but we certainly have to be able to discuss and bring wisdom to the conversation. On the lines of wisdom, we also need to be careful about the definition of terms. I think we can require companies to reveal their code, but you guys who know about theory of computing know that looking at code doesn't mean you have a clue what it is actually doing. There are fundamental questions that are still undecided. There is code obfuscation, and companies will exploit this if they don't want to reveal what they are doing. In the long run, I think we are going to have a situation where for example I am going to be represented by an artificial intelligence online which is going to go around and negotiate on my behalf, buy my airline tickets, etc. And this is another source I think of potential unfairness in the world. Take the artificial intelligence and replace it for example with a lawyer. People who can afford very good lawyers are going to win negotiations against people who can only afford much less expensive lawyers. And you can have a similar situation perhaps with artificial intelligences. The one that is going to represent me is perhaps not as good as the one that may represent a much richer person. And this is going to be exacerbated because things are going to happen really, really fast. So that's a whole another level of fairness that needs to be talked about.

At this point, Ricardo Baeza-Yates closed the panel discussion by thanking the panelists and the audience.

References

ACM US Public Policy Council. (2017). Statement on algorithmic transparency and accountability.

Abhinav Maurya is a graduate student at Carnegie Mellon University focusing on machine learning and public policy. He has previously studied at VJTI Bombay and IIT Bombay, and worked at Microsoft Redmond before heading to Pittsburgh to pursue doctoral studies.

AI Policy Matters

Larry Medsker (George Washington University; lrm@gwu.edu)
DOI: 10.1145/3236644.3236650

Abstract

AI Policy is a regular column in AI Matters featuring summaries and commentary based on postings that appear twice a month in the AI Matters blog (https://sigai.acm.org/aimatters/blog/). Selected posts are summarized in issues of *AI Matters*.

Introduction

The SIGAI Public Policy goals are to

- promote discussion of policies related to AI through posts in the AI Matters blog on the 1st and 15th of each month,
- help identify external groups with common interests in AI Public Policy,
- encourage SIGAI members to partner in policy initiatives with these organizations, and
- disseminate public policy ideas to the SIGAI membership through articles in the newsletter.

I welcome everyone to make blog comments so we can develop a rich knowledge base of information and ideas representing the SIGAI members.

EasyChair Data Privacy Issues

An emerging issue involves the data privacy of SIGAI and ACM members using EasyChair to submit articles for publication, including the AI Matters Newsletter. When trying to enter a new submission through EasyChair, the following message appears: AI Matters, 2014-present, is an ACM conference. The age and gender fields are added by ACM. By providing the information requested, you will help ACM to better understand where it stands in terms of diversity to be able to focus on areas of improvement. It is mandatory for the submitting author (but you can select "prefer not to submit") and it is desirable that you fill it out for all authors. This information will be deleted from EasyChair after the conference.

To evaluate the likelihood of privacy protection, one should pay attention to the EasyChair Terms of Service [1], particularly Section 6 Use of Personal Information. More investigation may allow us to decide the level of risk if our members choose to enter personal information.

Your Public Policy Officer is working with the other SIGAI officers to clarify the issues and make recommendations for possible changes in ACM policy. Please send your views on this issue to SIGAI and contribute comments to the SIGAI Blog.

AI Terminology Matters

In the daily news and social media, AI terminology is a part of the popular lexicon for better or for worse. AI technology is both praised and feared in different corners of the society. Big data practitioners and even educators add to the confusion by misusing AI terms and concepts.

"Algorithm" and "machine learning" may be the most prevalent terms that are picked up in popular dialogue, including in the important fields of ethics and policy. The ACM and SIGAI could have a critical educational role in the public sphere. In the area of policy, the correct use of AI terms and concepts is important for establishing credibility with the scientific community and for creating policies that address the real problems.

In recent weeks, interesting articles have appeared, authored by writers of diverse degrees of scientific expertise. A June issue of The Atlantic has an article [2] by Henry Kissinger entitled "How the Enlightenment

[1] https://easychair.org/terms.cgi

[2] https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/

Ends" with the thesis that society is not prepared for AI. While some of the understanding of AI concepts can be questioned, the conclusion is reasonable: "AI developers, as inexperienced in politics and philosophy as I am in technology, should ask themselves some of the questions I have raised here in order to build answers into their engineering efforts. The U.S. government should consider a presidential commission of eminent thinkers to help develop a national vision. This much is certain: If we do not start this effort soon, before long we shall discover that we started too late."

In May, The Atlantic had an article about the other extreme of scientific knowledge by Kevin Hartnett entitled "How a Pioneer of Machine Learning Became One of Its Sharpest Critics" [3]. He writes about an interview with Judea Pearl about his current thinking, with Dana Mackenzie, in *The Book of Why: The New Science of Cause and Effect* [4]. The interview includes a criticism of deep learning research and the need for a more fundamental approach.

Focusing back on policy, I recently attended a DC event of the Center for Data Innovation (https://www.datainnovation.org/) on a proposed policy framework to create accountability in the use of algorithms. They have a report [5] on the same topic. The event was another reminder of the diverse groups in dialogue, in the public sphere, on critical issues in AI and the need to bring together the policymakers and the scientific community. SIGAI can have a big role to play.

Potential Revival of the OTA

As a small agency within the Legislative Branch, the Office of Technology Assessment (OTA) originally provided the United States Congress with expert analyses of new technologies related to public policy. But OTA was defunded and thereby ceased operations in 1995. A non-binding Resolution was introduced in the House of Representatives last week by Reps. Bill Foster (D-IL) and Bob Takano (D-CA) (press release), and Sen. Ron Wyden (D-OR). It is expected to introduce a parallel bill in the Senate, expressing the non-binding "sense of Congress" that the agency and its funding should be revived. New coordinated efforts are now underway as well, among many groups, to urge Congress to do exactly that.

Our colleagues at USACM have delivered letters of support for an inquiry into whether restoring OTA or its functions to the Legislative Branch would be advisable to the leaders of the House and Senate Appropriations Committees. The House Subcommittee met recently and voted to advance legislation to fund the Legislative Branch for FY 2019 to the full House Appropriations Committee but without addressing this issue. The full Committee's meeting, at which an amendment to provide pilot funding for an inquiry into OTA-like services, is expected later in May. The Senate's parallel Subcommittee and full Appropriations Committee is expected to act later this spring or early summer on the Legislative Branch's FY19 funding bill. OTA-related amendments could be offered at either of their related business meetings. See the letter [6] from USACM to leaders in the House and Senate Appropriations Committees

White House AI Summit on AI for American Industry

From the report[7]:

On May 10, 2018, the White House hosted the Artificial Intelligence for American Industry summit, to discuss the promise of AI and the policies we will need to realize that promise for the American people and maintain U.S. leadership in the age of artificial intelligence. "Artificial intelligence holds tremendous potential as a tool to empower the American worker, drive growth in American industry, and improve the lives of the American people. Our free market approach to scientific discovery

[3] https://www.theatlantic.com/technology/archive/2018/05/machine-learning-is-stuck-on-asking-why/560675/

[4] https://www.basicbooks.com/titles/judea-pearl/the-book-of-why/9780465097609/

[5] https://www.datainnovation.org/category/publications/reports/

[6] https://www.acm.org/binaries/content/assets/public-policy/usacm/2018-usacm-letter-ota-funding.pdf

[7] https://www.whitehouse.gov/wp-content/uploads/2018/05/Summary-Report-of-White-House-AI-Summit.pdf

harnesses the combined strengths of government, industry, and academia, and uniquely positions us to leverage this technology for the betterment of our great nation." - Michael Kratsios, Deputy Assistant to the President for Technology Policy

The summit brought together over 100 senior government officials, technical experts from top academic institutions, heads of industrial research labs, and American business leaders who are adopting AI technologies to benefit their customers, workers, and shareholders.

Issues addressed at the 2018 summit are as follows:

- Support for the national AI R&D ecosystem "free market approach to scientific discovery that harnesses the combined strengths of government, industry, and academia."
- American workforce that can take full advantage of the benefits of AI "new types of jobs and demand for new technical skills across industries . . . efforts to prepare America for the jobs of the future, from a renewed focus on STEM education throughout childhood and beyond, to technical apprenticeships, re-skilling, and lifelong learning programs to better match America's skills with the needs of industry."
- Barriers to AI innovation in the United States included – "need to promote awareness of AI so that the public can better understand how these technologies work and how they can benefit our daily lives."
- High-impact, sector-specific applications of AI - "novel ways industry leaders are using AI technologies to empower the American workforce, grow their businesses, and better serve their customers."

Bias, Facebook, and Google

Current events involving FaceBook and the use of data they collect and analyze relate to issues addressed by SIGAI and USACM working groups on algorithmic accountability, transparency, and bias. The players in this area of ethics and policy include those who are unaware of the issues and ones who intentionally use methods and systems with bias to achieve organizational goals. The issues around use of customer data in ways that are not transparent, or are difficult to discover, not only have a negative impact on individuals and society, but are also difficult to address because they are integral to the business models upon which such companies are based.

A recent Forbes article[8] "Google's DeepMind Has An Idea For Stopping Biased AI" discusses research that addresses AI systems that spread prejudices that humans have about race and gender – the issue that biased decisions may be made by artificial intelligence systems when trained on biased data. An example cited in the article include facial recognition systems that have been shown to have difficulty in properly recognizing black women.

Machine-learning software is rapidly becoming widely accessible to developers across the world, many of whom are not aware of the dangers of using data containing biases. The Forbes piece discusses an article "Path-Specific Counterfactual Fairness," [9] by DeepMind researchers Silvia Chiappa and Thomas Gillam. Counter-factual fairness refers to methods of decision-making for machines and ways that fairness might automatically be determined. DeepMind has a new division, DeepMind Ethics & Society, that addresses this and other issues on the ethical and social impacts of AI technology.

The Forbes article quotes Kriti Sharma, a consultant in artificial intelligence with Sage, the British enterprise software company, as follows: "Understanding the risk of bias in AI is not a problem that technologists can solve in a vacuum. We need collaboration between experts in anthropology, law, policy makers, business leaders to address the questions emerging technology will continue to ask of us. It is exciting to see increased academic research activity in AI fairness and accountability over the last 18 months, but in truth we aren't seeing enough business leaders, companies applying AI, those who will eventually make AI mainstream in every aspect of our lives, take the same level of responsibility to create unbiased AI."

[8] https://www.forbes.com/sites/parmyolson/2018/03/13/google-deepmind-ai-machine-learning-bias/#43851fb26829

[9] https://deepmind.com/applied/deepmind-ethics-society/

News Matters

- The statement of the European Group on Ethics in Science and New Technologies on “Artificial Intelligence, Robotics and ‘Autonomous’ Systems,” was published March 9: http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf. The statement calls for the EC to “launch a process that paves the way towards a common, internationally recognized ethical and legal framework for the design, production, use and governance of artificial intelligence, robotics, and ‘autonomous’ systems.”
- President Donald Trump recently tapped Obama-era deputy U.S. CTO Ed Felten to serve on the Privacy and Civil Liberties Oversight Board https://www.pclob.gov/
- AAAS Forum on Science & Technology Policy, Washington, D.C., June 21 22, 2018. https://www.aaas.org/page/forum-science-technology-policy?et_rid=35075781&et_cid=1876236.

Larry Medsker is Research Professor of Physics and a faculty member in the Data Science graduate program at The George Washington University. Dr. Medsker is a former Dean of the Siena College School of Science, and Professor in Computer Science and in Physics, where he was a co-founder of the Siena Institute for Artificial Intelligence. His research and teaching continues at GW on the nature of humans and machines and the impacts of AI on society and policy[a]. Professor Medskers research in AI includes work on artificial neural networks and hybrid intelligent systems. He is the Public Policy Officer for the ACM SIGAI.

[a] http://humac-web.org/

The New Robber Barons: The Optimistic Objectivists Overturning the Old Order

Toby Walsh (UNSW Sydney; tw@cse.unsw.edu.au)
DOI: 10.1145/3236644.3236651

The Wrong Stuff?

You hear a lot today about Artificial Intelligence and Ethics. For good reason. As algorithms get smarter, they are increasingly making decisions that impact on people's lives. And so we need to be careful that these algorithms are fair and transparent. But it's not only the algorithms that concern me.

If Tom Wolfe hadn't died recently, he might be turning his critical pen towards people like Mark Zuckerberg and Travis Kalanick, playfully describing the disruption that their companies bring to our lives on a daily basis.

Just as his iconic novel, "The Bonfire of the Vanities" focused on the financiers whose greed defined the economy of the 1980s, Wolfe might today focus on the behaviour of these entrepreneurs whose technological innovations are overthrowing the old economy, creating entirely new digital marketplaces.

And rather than the greed of the 1980s, ethics might be at the centre of Wolfe's attention today. Not the ethics of the algorithms running these business. For algorithms don't have ethics. Even smart algorithms don't have ethics.

Algorithms are just bits of mathematics. Algorithms do, however, capture the ethics of the people behind them. And there is so much material Wolfe could write about this in 2018.

Wind back the clock nearly two years. In October 2016, the investigative nonprofit newsroom ProPublic discovered that Facebook let advertisers exclude black, Hispanic, and other "ethnic affinities" from seeing adverts.

In the United States, housing and job adverts that exclude people based on race, gender and similar factors are prohibited by the Fair Housing Act of 1968 and the Civil Rights Act of 1964. Facebook admitted this was "a failure" and promised to prevent such discrimination in the future.

Over one year later, in November 2017, ProPublica found that Facebook was still allowing such adverts to be placed. This is not a failure to write ethical algorithms. This is a failure to care.

Give any programmer access to the Facebook code base, and it would take less than an afternoon to remove such functionality from the system. In March 2018, fair housing groups filed a lawsuit in the federal courts to stop Facebook selling discriminative adverts. Perhaps this will be enough to make Facebook care?

Children at risk

It is easy to pick on almost every other large technology company. Take Google for example. In the United States, the Children's Online Privacy Protection Act (COPPA) prohibits operators of websites from collecting any personal information on children without parental consent.

In April 2018, twenty advocacy, consumer and privacy groups filed a complaint to the Federal Trade Commission that Google is violating COPPA by collecting information without parental consent about the location, phone numbers and viewing habits of children watching YouTube videos.

Google's lawyers have responded that YouTube is not intended for children aged less than 13 as its terms of service require users to be over this age. This is just laughing in our faces.

The managers of YouTube know that millions of children are watching its videos. Search for "kids" videos on YouTube and you get over 177 million results. Or go check out the official Sesame Street, National Geographic Kids or Peppa Pig channels on YouTube. These clearly aren't aimed at adults.

Google is using the information it collects on YouTube to sell adverts targeted at children.

Google doesn't want to fix this because they make a lot of money out of having children watch YouTube and collecting information on them.

Incidentally Mark Zuckerberg has gone on record saying that he intends to fight COPPA so he can get kids on Facebook at an earlier age. This is despite the concerns of child development experts about the dangers of exposing children to social media at a young age.

The New Robber Barons

Now this isn't the first time we've faced such problems. In the first Industrial Revolution, some of the first to benefit became known as the "Robber Barons".

And chief amongst the Robber Barons was the industrialist John D. Rockefeller, arguably the wealthiest man ever to live in modern times. Rockefeller was notorious for his unethical and illegal business practices that helped his company, Standard Oil control up to 90 percent of the world"s oil refineries.

The History of the Standard Oil Company, published by Ida Tarbell in 1904, famously described the company's espionage, price wars and courtroom antics that allowed it dominate the oil business. Eventually Standard Oil became so powerful that it had to be broken up into 34 new companies.

To balance the power of corporations, we also introduced institutions like unions, labour laws and the welfare state so that workers would share the benefits that industrialisation brought.

Today, we have a new set of Robber Barons, running digital monopolies and again receiving excessive benefits from the disruption brought about by new technology. History tells us that we will need to regulate their monopolies, perhaps even break them up. In addition, history tells us that we may have to build new institutions to ensure that the new technologies improve the common good.

We may also need to reform the modern corporation – let's not forget that the corporation was also an invention of the first industrial revolution – as well as strengthen worker and digital rights. And finally we may need to change how we tax corporations and the individual so we share better the benefits.

We did it once. And now it seems, we need to do it a second time,.

Acknowledgments

The author acknowledges support from the European Research Council (AMPLify Advanced Grant 670077) and the Asian Office of Aerospace Research and Development (Grant FA2386-15-1-4016).

Toby Walsh is Scientia Professor of Artificial Intelligence at the University of New South Wales, Group Leader at Data61 and Guest Professor at the Technical University of Berlin. He is a strong advocate for limits to ensure AI is used to improve our lives. He has been a leading voice in the discussion about lethal autonomous weapons (aka killer robots) speaking at the UN in New York and Geneva. He is a Fellow of the Australia Academy of Science and of the Association for the Advancement of Artificial Intelligence. He has authored two books on AI for a general audience.

The Challenges of Outreaching

Jinhong K. Guo, ACM SIGAI membership and outreach officer (Lockheed Martin Advanced Technology Laboratories ; jinhong.guo@lmco.com)

Rosemary Paradis, ACM SIGAI Secretary/Treasurer (Leidos Data Analytics Products and Services; rosemary.paradis@leidos.com)

DOI: 10.1145/3236644.3236652

Introduction

According to Wikipedia, *"outreaching is an activity of providing services to any populations who might not otherwise have access to those services".* Most people would have been the target of a fund raising campaign at some point, be it public radio during its annual fund raising or an alma mater fund raising event. While some of us may respond because it is a good cause or it aligns with our interests, outreaching has always been very challenging. We recently met with some ACM groups that have been successful in getting members and outreaching to new people. This short paper summarizes what we learned from the conversations and discovered from our recent activities.

Communications and Activities to Reach a Wider Audience

The goal of our special interest group, SIGAI, is to promote and support the growth and application of AI principles and techniques. With the current research and application interests in AI technology, now is a great time to garner people's interests in joining SIGAI. However, with many similar professional societies, such as AAAI, spreading "our brand" to a wide audience convincingly still has to rely on frequent communications. Northeast Ohio ACM group (NEO) is one of the most active local groups in ACM. Their key to success is engagement through local gathering, especially the new trends of incubators engaged in engineering and citizen science, such as Hackspaces and Makerspaces. Their leadership frequently looks though local list of meetups to find interesting meetings to attend.

Motivation is essential to all actions. A big incentive for member of SIGCHI is the CHI conference which requires a membership with submission. They also offer reduced registration at conferences sponsored and co-sponsored by SIGCHI. Additionally, they cooperate with related societies, such as the International Federation for Information Processing and the Usability Experience Professionals' Association.

SIGAI Activities to Meet the Challenge

Outreach is not only about establishing links, especially just temporary links. Outreach depends on building long-lasting relationships that are mutually beneficial. The SIGAI webinar series is a good example. The industry liaison committee holds webinars every month, with the topics selected for the interests to industry practitioners and academics alike. The initial intention of running a fall and spring series turned into an ongoing event; these webinars are well attended with interested technologists from around the world.

This year, we took a survey of current and past SIGAI members to find out what they would be interested in as members of SIGAI. This survey indicates that people are interested in the Webinars and conferences. SIGAI sponsored 1st AAAI/ACM conference on Artificial Intelligence, Ethics, and Society is a well-received conference. If there are other areas of interest that you would be interested in, please contact the authors of this article.

It is important to find the right people to reach with the niche that is unique to SIGAI. The AAAI is already established as the premier forum for AI research. The SIGAI Webinars and conference on AI and ethics and society show an area that SIGAI can still make a difference with focus on industry and applications. Our niche may just be the applications of AI that are now flooding the workspace. What are they, how are they built and how are they being built and implemented.

People join SIGAI for the networking opportunities and resources it brings. The most common reasons for those who did not renew is that they did not take advantage of the activities and benefits. It is a challenge to us to provide members with activities and benefits that they are interested in. The recent increase in membership shows that we are going in the right direction, especially with the interests shown in the Webinars and sponsored conferences. Let us keep it up!

Acknowledgments

We are grateful to Tracey Hughes and Cameron Hughes from NEO ACM for the time spent taking with us and providing us with suggestion on how to keep members engaged in activities. We are grateful to the SIGCHI officers, especially Vice President for Membership and Communication, Regina Bernhaupt and her group of officers for sharing with us their knowledge on membership outreach and retaining. We also would like to thank the ACM SIG coordinators, especially Sunita Jaswal for the information on the chapters program and benefits.

Dr. Jinhong K. Guo is a senior researcher at Lockheed Martin Advanced Technology Laboratories. Her research interests are primarily in AI and Applications, including Machine Learning, NLP and Distributed Intelligent Agents

Dr. Rosemary Paradis is a Principal Machine Learning Research Scientist for Leidos, Data Analytics Products and Services, where she builds machine learning models and data analytics for a number of Big Data applications. Her work focuses on data and text analytics, computational linguistics and natural language processing.

Is There an AI Cosmology?

Cameron Hughes (Northeast Ohio ACM Chair; cameronhughes@acm.org)
Tracey Hughes (Northeast Ohio ACM Secretary; tracey.hughes@neoacmchapter.org)
DOI: 10.1145/3236644.3236653

Is There an AI Cosmology?

In the past fifteen years artificial intelligence has changed from being the preoccupation of a handful of scientists to a thriving enterprise that has captured the imagination of world leaders and ordinary citizens alike.

This is a quote taken from the article "Artificial Intelligence in Transition" written by Peter E. Hart published in AI Magazine Volume 5 Number 3 over 30 years ago in 1984. The first paragraph of the article states:

> THE FIELD OF ARTIFICIAL INTELLIGENCE is in the midst of a deep and irreversible structure change. The older research institutions that were almost alone on the AI landscape (at least in North America) in the late 60s and early 70s have been joined by a host of newer ones; new products based on the fruits of AI research have begun to appear and the public at large is beginning to believe that "intelligence" can be put in machines.

Something about that paragraph has a familiar ring. AI Magazine was not alone in helping to clarify the status and progress of Artificial Intelligence. Two years earlier, the ACM's SIGART Bulletin, a somewhat progenitor of our own AI-Matters newsletter, published in Issue 79, January 1982 an article entitled "DIRECTIONS FOR AI IN THE EIGHTIES". The article starts off with the following paragraph:

> The 1980s will be an exciting decade for artificial intelligence, one in which we can expect to see considerable progress on scientific questions, widespread application of AI techniques to programs of practical significance, and major changes in the social political and economic structure of the field. I will attempt to identify some of the commercial, scientific and sociological considerations that I think will be important in coming years. In particular, I will introduce the metaphor of deep versus surface systems, and pose some challenges arising from the recognition of this distinction.

The article goes on to explain some of the short comings of AI efforts in the 70s and why things will be different this time for the AI technologies and techniques of the 80s. Both articles are contextualized at a time when there was an intense fervor, anticipation and excitement for all things AI. At the time there was incredible hype for latest AI technology. The technical media and many business journals were virtually intoxicated with AI's flagship technology, "Expert Systems," which were everywhere and being touted as the smart panaceas for almost everything from sports predictions to some of societies largest challenges. Over 30 years later, the AI community, the tech media, business journals, and pop culture no longer have the zeal for, or infatuation with the promise of endless applications of the "Expert System".

But if we could indulge ourselves (just for a moment) in a little time travel, we would see striking parallels between the frenzy over Artificial Intelligence's Expert Systems Era and the current frenzy over Artificial Intelligence's Machine Learning Era. The listing shows the domain and range of some of the more notable expert systems from the 70s and 80s and some of the current applications of machine learning. Paul Harmon and David King in their landmark book "Expert Systems Artificial Intelligence In Business" published in 1985 portray AI's transformation of business and industry at that time as being imminent. The inevitable magnanimous impact of each of the expert systems from the listing is detailed in Harmon and King's book. The impending transformation of our society by AI technologies in the form of Expert Systems was over 30 years ago.

70s - 80s EXPERT SYSTEMS

- **MYCIN** consulted and advised on meningitis.
- **DENDRAL** performed spectroscopic analysis on unknown molecules.
- **HEARSAY I and II** performed signal and speech processing.
- **PUFF** interpreted measurements from respiratory tests administered to patients in pulmonary lung function laboratory.
- **GENESIS** consulted on the nature of DNA molecules.
- **Drilling Advisor** assist Oil rig supervisors in resolving problems relating to the drilling process.

CURRENT MACHINE LEARNING APPLICATIONS

- **InnerEye** helps radiologists identify and analyze 3-D images of cancerous tumors.
- **FingerID** learns to predict molecular fingerprints from a large collection of MS/MS spectra then retrieves and rank candidate molecules from a given large molecular database.
- **Apple's ResearchKit** is used in the treatment of Parkinson's disease and Asperger's Syndrome by allowing users to access facial recognition apps that assess their conditions over time; their use of the app feeds ongoing progress data into an anonymous pool for future study.
- **Chatbots** speech recognition capability is used to identify patterns in patient symptoms to form a potential diagnosis, prevent disease and/or recommend an appropriate course of action.
- **Deep Learning** is used to train algorithms to recognize cancerous tissue at a level comparable to trained physicians.
- **Machine Learning** used to help optimize drilling operations and facility/well management as well as investment decisions in exploration and production.

It would seem the quest for the promises of AI undergoes these periodic hype-frenzy-transformation-adaptation cycles. It appears we are riding the crest of one of these cycles now with the machine learning + big data phenomenon. There was certainly one in the 80s. There was one in the 40s even though it did not have the designation "Artificial Intelligence". In the 40s and 50s similar goals were pursued under the auspices of Cybernetics but that's the subject for another article.

Fast forward to today, and we have pronouncements by prominent researchers, venture capitalists, futurists, government agencies, and some multinational corporations about the impending transformation of our society at the hands of AI technologies. It is a fact that Expert Systems were transformative but for all their applications they didnt fulfill the promise AI. In fact in 2018, the fervor, hype and frenzy that Expert Systems brought society has all but been forgotten. It's almost as though AI is only just now starting to return dividends. For example on May 3, 2016, Ed Felton who was Deputy U. S. Chief Technology Officer in his "Preparing for the Future of Artificial Intelligence" letter wrote:

> There is a lot of excitement about Artificial Intelligence (AI) and how to create computers capable of intelligent behavior. After years of steady but slow progress on making computers "smarter" at everyday tasks, a series of breakthrough in the research community and industry have recently spurred momentum and investment in the development of this field.

For the complete letter, see:

obamawhitehouse.archives.gov/blog/2016/05/03/preparing-future-artificial-intelligence

The letter goes on to talk about the impending impact of AI on the medical field and other areas. It introduces an entirely new set of enthusiasms for AI. The consortium Partnership on AI formed in 2016 consisting of a wide range of companies and organizations such as Amazon, Facebook, IBM, Apple, Google, Deep Mind, AAAI, ACLU, ACM, and Microsoft has announced:

"We are at an inflection point in the development and application of AI technologies. The upswing in AI competencies, fueled by data, computation, and advances in algorithms for machine learning, perception, planning, and natural language, promise great value to people and society". With organizations such as

these behind this kind of statement one is compelled to believe that very significant benefits are in the cards for people and society at the hands of AI technologies. The 100 year study on Artificial Intelligence [1]delves in great detail about the status and potential promise of AI. The AI 100 report is informative. But so was so many of the reports and reporting from the Expert System era of the 1980s. How do we separate fact from fiction? How do we separate promising results from the hype that so often follows the publishing of those results? What is it about research into Artificial Intelligence that seems to lead to an almost periodic cycle of promising results, hype, frenzy, and then adaptation?

Simon Natale and Andrea Ballatore in their paper "Imagining the thinking machine: Technological Myths and the Rise of Artificial Intelligence" posits the notion that the rise of Artificial Intelligence was accompanied by a powerful creation myth:

"the creation of a thinking machine, which would be able to perfectly simulate the cognitive faculties of the human mind ..."

Natale and Ballatore go on to challenge and expose many of the phenomena associated with the rise then fall and then rise again of various incantations of Artificial Intelligence. We strongly recommend a thorough read of Natale and Ballatores work for anyone that may want a more sober or different interpretation of what's going on with this AI stuff.

So what is a dutiful researcher, practitioner, student or interested onlooker supposed to make of the AI landscape? How should they interpret the hype-frenzy cycles, the conflicting information and definitions, the legitimate results, the exaggerated results, the real impacts, the promising futures, the feared and dreaded disappointing futures, the thousands of disparate blogs and advertisements all simultaneously claiming to refer to Artificial Intelligence?

Is There a Need for an AI Cosmology?

It was these considerations (and a few more) that lead of to the question of whether there is a true Cosmology for Artificial Intelligence and whether one is needed. Typically, the term 'Cosmology' is used in conjunction with the study of the Universe. One convenient definition for Cosmology "is the study, origin, evolution and the eventual fate of the Universe". It is important to note that there are various theological, mythical, and scientific cosmologies for our Universe. But could we apply the concept to the notion of "An Artificial Intelligence Universe"? Is there a cosmology of Artificial Intelligence? Can we talk about the study, origin, evolution and the eventual fate of Artificial Intelligence? If there is a cosmology for AI, what would it look like? If there are multiple cosmologies for AI, what would be the justification for more than one? How would they differ?

Any valid AI Cosmology would force us to define its origin. And in defining its origin, we would have to be very specific in uncovering the goals of AI and the meaning of the term "Artificial Intelligence". Does it refer human reasoning, memory, imagination, creativity, or other cognitive, or neurological capabilities? What is it's structure? How is it evolving? What is it's fate? Is AI an intersection of other areas of study? For example, Figure 1 shows a simplistic venn diagram of what might represent what we call AI.

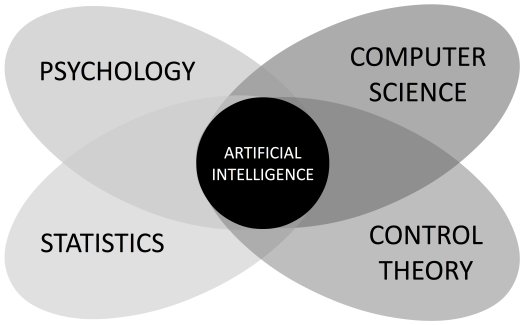


Figure 1: Venn diagram of AI.

We would be able to examine the major structures and processes at work in its evolution. An AI Cosmology would allow us to predict the vector of its fate. An AI Cosmology would depend on taxonomy. Figure 2 is a very simply but interesting starting point for a discussion on AI taxonomy.

A taxonomy for AI will help on the development of an ontology for AI or vice versa. A valid AI Cosmology would ultimately give us a discrete story for AI having a beginning, mid-

[1] ai100.stanford.edu/2016-report

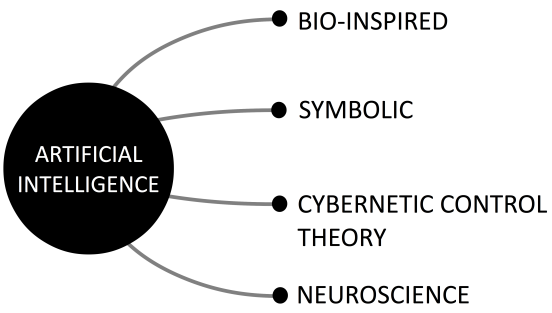


Figure 2: Stem of an AI taxonomy.

dle and end, allowing us to:

- Evaluate hype-frenzy cycles and predictions concerning AI
- Contextualize AI innovations
- Inform our moral and ethical discussion

But first things first. Just as our Big Bang Cosmology attempts to give us the true picture of the origins of the Universe, an AI Cosmology should attempt to give us a true picture of the origins of Artificial Intelligence.

We would like to dedicate a column in AI Matters to the discussion and possible developmental beginnings for an AI Cosmology, beginning with identifying the true origins of Artificial Intelligence.

Cameron Hughes is a computer and robot programmer. He is a Software Epistemologist at Ctest Laboratories where he is currently working on A.I.M. (Alternative Intelligence for Machines) and A.I.R (Alternative Intelligence for Robots) technologies. Cameron is the lead AI Engineer for the Knowledge Group at Advanced Software Construction Inc. He is a member of the advisory board for the NREF (National Robotics Education Foundation) and the Oak Hill Robotics Makerspace. He is the project leader of the technical team for the NEOACM CSI/CLUE Robotics Challenge and regularly organizes and directs robot programming workshops for varying robot platforms. Cameron Hughes is the co-author of many books and blogs on software development and Artificial Intelligence.

Tracey Hughes is a software and epistemic visualization engineer at Ctest Laboratories. She is the lead designer for the MIND, TAMI, and NO-FAQS projects that utilize epistemic visualization. Tracey is also a member of the advisory board for the NREF (National Robotics Education Foundation) and the Oak Hill Robotics Makerspace. She is the lead researcher of the technical team for the NEOACM CSI/CLUE Robotics Challenge. Tracey Hughes is the co-author with Cameron Hughes of many books on software development and Artificial Intelligence.

Non-intervention policy for autonomous cars in a trolley dilemma scenario

Bianca Helena Ximenes (Informatics Department, Universidade Federal de Pernambuco; bhxmm@cin.ufpe.br)

DOI: 10.1145/3236644.3236654

Introduction

This column presents early work on human biases and preferences described in research from Philosophical and Social Sciences, and discusses their impact in AI ethics, especially concerning autonomous vehicles.

Human nature and decision biases

In many contexts, humans prefer to withdraw their participation in scenarios where decision is too complex or have convoluted ethical implications. Often, they will let chance, or time, or another exogenous factor force a decision, so they are spared the choice and its consequences. Surely, that effectively means making a choice in the end; but not opting explicitly makes it easier for humans to deal with their own conscience.

By recognizing such effects, the present column aims to discuss the following problem: *Is a non-intervention policy in trolley dilemma scenarios a desirable way for humans to interact with autonomous vehicles?*

The discussion that follows is based on the premise that practitioners responsible for autonomous vehicles have a moral obligation of ensuring their full functionality to the best of their ability, and that saving lives is a golden rule. However, it is also based on the premise that scenarios such as that of the trolley dilemma will unfortunately be present and understanding human limitations and preferences might prove useful to modeling.

Killing or letting die?

One of the paramount aspects of trying to establish a rank of societal priorities from human research or questionnaires is the framing effect. Seminal research, such as the one that granted the Nobel of Economic Science to Daniel Kahneman and Amos Tversky shows how the same two sets of options, framed differently, yielded a completely different final collective preference concerning what to do in a critical scenario of an epidemic. Moreover, one of the factors that drove the change in the volunteers opinion was the use of the word kill, which is negatively charged and directly related to trolley dilemma scenarios (Tversky & Kahneman, 1974) (Tversky & Kahneman, 1981).

One of the earliest discussions of the trolley dilemma itself had Thomson argue that some people found that letting die had a different moral weight than killing, and different choices ensued according to such perception (Thomson, 1976). She followed presenting multiple dilemmas, each framed with a slight difference from previous others, and the results of what was deemed admissible or otherwise varied according to each factor. Often these factors were information of how the scenario came to be, and relied on extensive background information, such as the person in the tracks was a child, the person was there illegally and knowingly, the person was put there by a villain, or the person was randomly assigned to be there. These scenarios brought other papers that discussed each nuance more detailedly. This is summarized in Figure 1.

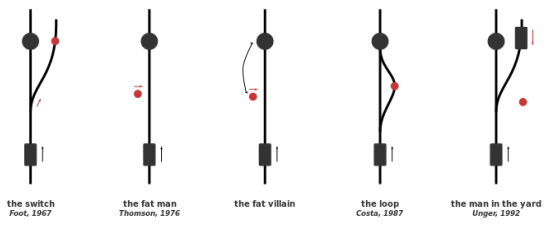


Figure 1: Trolley dilemma scenarios of Foot (1967), Thomson (1976), Costa (1987), and Unger (1992)

None of these nuances will be taken into account by an autonomous vehicle, as such information will not be available at the time of an accident. Therefore, it is not possible to rely on human objectivity when establishing an immutable ranking as a parameter to an algorithm.

Finally, in the more recent article (Waldmann & Dieterich, 2007), authors reframe once again the trolley problem and propose a simulation experiment. They find objectively that the moral standards fluctuate according to different framing of scenarios and background information, and classify it as an intervention myopia. Hence, non-intervention might be a more viable path.

A universal guideline for saving lives

The MIT Moral Machine (*MIT Moral Machine*, 2016) makes it clear that humans value actions and lives differently. Figure 2 depicts the general preference on whose life or which principle should be considered as most important when deciding whether and where to divert a cars course.

Not all preferences were rated and ordered in a single rank, being instead translated into preferences between pairs of factors that could be demographical (i.e.: age, fitness, gender, species) or more related to personal belief (i.e.: avoiding intervention, individuals social value, number of lives, protecting passengers, upholding traffic laws).

By using millions of data points to train a Machine Learning model in sufficient variable scenarios, it would be possible to achieve a general complete ranking of the value of lives and actions that reflected the judgments of the majority, use it to establish rules for autonomous cars, and release them on the streets once they are ready.

Doing so is technically possible. However, choosing who to let die might be too serious a choice to be left up to personal opinion or to an algorithm, especially when it is universally applicable across different countries and cultures. It can be argued that it is neither fair nor admissible to extend the opinions of the majority to an issue of literal life or death; neither would it be ethical to rank humans' lives, especially when there are no apparent reasons

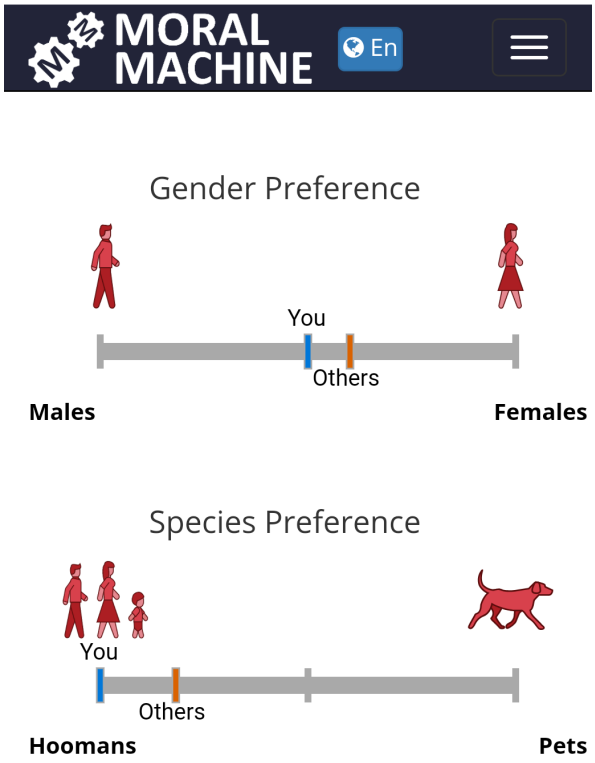


Figure 2: Partial screenshot of the result of the MIT Moral Machine judgment game. The results highlight how the judgment of the taker compares to the overall results

for specific choices other than personal preference. This scenario of automated pre-defined choice could be considered a breach of Human Rights, as one specific demographic profile would be in practice marked as inferior or less socially costly if they were to die. Considering the recent UK House of Lords' document on Artificial Intelligence and its several related subjects and implications. It states in one of its summary points that "The autonomous power to hurt, destroy or deceive human beings should never be vested in artificial intelligence." However, by implicitly attributing value to different lives and personal characteristics, the result could be framed as the AI ultimately choosing to hurt or destroy a human being - the one who ranks lower according to the algorithm. This choice is not to be made deliberately, as it is one step further weaponizing AI to act against specific groups of people.

Another issue that comes up from Social Science research is the impossibility of carrying

out perfectly democratic votes that conform to the principles of being transitive, reflexive, and complete when there are more than two choices. This is discussed in the Impossibility Theorem by Kenneth Arrow and underpins much of the Economic Theory of choice. Arrow even concludes that the only fair system after all would be a dictatorship, even though it was not desirable (Arrow Kenneth, 1951). Arrows theorem has been addressed in research throughout the years, but it is not a solved conundrum (Frohock, 1980). As seen in the scenarios of the Moral Machine, there are well over 2 options one can choose from, which falls into a classical scenario where a democratic, fair choice is not attainable.

Choosing not to choose

The difficulty of making choices is also studied in other behavioral sciences, and go beyond cognitive biases of the human incapability of calculating outcomes into the field where humans attempt to avoid explicit choices altogether, preferring to let time run out than commit to a single option (Shin & Ariely, 2004)

Leaving the choice of who to let die to chance "or non-interventionism" may carry other results that have to be further researched but show promise. It unburdens the autonomous vehicle user both because they know the car will not explicitly choose another life over their own (in the case the algorithm is set to save pedestrians), and because they do not need to feel responsible for complying with an algorithm that ranks and weighs people's lives. The research of Kelly shows that parents who have experienced moral decisions derived from first pregnancies where the fetus had disabilities or genetic deficiencies detected while in the womb only chose to try to conceive again 34.5% of the time. On the other hand, parents who experienced first pregnancies where the fetus died for reasons other than malformation were much more prone to trying to conceive again, with over 85% opting for parenthood. Kelly concluded that the families in the former case not only had an emotional burden to carry, they also had to make ethically complex choices, such as terminating the pregnancy or deciding whether a disability was indeed good reason for an abortion. In the end, the majority of them preferred avoiding the issue altogether.

The direct repercussion for implementing a self-driving algorithm that does not take into consideration the unwillingness of humans to be faced with complex, ambiguous choices, is that autonomous cars may face resistance in adoption. The benefits of self-driving cars are such that they are being embraced and regulated by governments around the world, seen as a way to move past the most ubiquitous reason of traffic accidents and deaths: human error. An example is the SELF-DRIVE Act passed in 2017 by the United States of America House of Representatives (The Senate and House of Representatives of the United States of America, 2017). But that reason by itself might not be enough for persuading users to make the change; especially if they find that they do not agree with ranking parameters and model results, or that they will face a complex moral conundrum every time they take the car out for a drive. It is one thing to be faced with an unwanted scenario, such as a trolley problem, and make a decision in the moment. It is another to leave the house knowing which decisions have been made.

Conclusions

Non-intervention as a policy for autonomous vehicles is something hard to discuss, but it may prove a viable option due to three major questions discussed in this column: (i) the framing effect and intervention myopia, due to humans being very sensitive to changes in context; (ii) the impossibility to reach with a universal rank across nations and cultures, and how model results can be more easily diverted to uses that were unintended; and (iii) the difficulty to deal with complex moral questions when the output is known or considered too risky.

By considering these topics, we can open new frontiers on moral, ethical, and AI research. It may be hard to accept humans are not able to control these scenarios, but open discussions need to be carried out to assess whether the benefits of a non-biased system a system that relies on chance can outweigh the perils that come with what humans are building.

References

Arrow Kenneth, J. (1951). Social choice and individual values. *Cowles Foundation*.

Frohock, F. M. (1980). Rationality, morality, and impossibility theorems. *American Political Science Review*, *74*(2), 373–384.

Kelly, S. E. (2009). Choosing not to choose: reproductive responses of parents of children with genetic conditions or impairments. *Sociology of Health & Illness*, *31*(1), 81–97.

Mit moral machine. (2016). http://moralmachine.mit.edu. (Accessed: 2018-05-30)

Shin, J., & Ariely, D. (2004). Keeping doors open: The effect of unavailability on incentives to keep options viable. *Management Science*, *50*(5), 575–586.

The Senate and House of Representatives of the United States of America. (2017). *Safely ensuring lives future deployment and research in vehicle evolution act, or self drive act (hr 3388).* (https://www.congress.gov/bill/115th-congress/house-bill/3388/text)

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist*, *59*(2), 204–217.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *science*, *211*(4481), 453–458.

Waldmann, M. R., & Dieterich, J. H. (2007). Throwing a bomb on a person versus throwing a person on a bomb: Intervention myopia in moral intuitions. *Psychological science*, *18*(3), 247–253.

Bianca Helena Ximenes
Bianca is a cross-disciplinary researcher and Doctoral Student at Universidade Federal de Pernambuco. Her research spans Human-machine boundaries and interaction, Innovation Management, Behavioral Economics, and Digital Ethics. Her community impact earned her the Google Developer Expert in Product Strategy recognition. She is member of the IEEE P7010 Working Group and a Sci-Fi aficionada.

Filtering and Planning for Resource-Constrained Mobile Robots

Tauhidul Alam (Florida International University; talam005@fiu.edu)

DOI: 10.1145/3236644.3236655

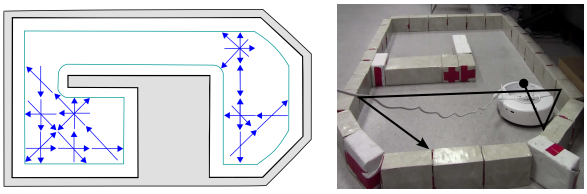
Note: This article is a summary of Tauhidul Alam's full dissertation that was submitted in partial fulfillment of the requirements to obtain Ph.D. degree in Computer Science, and defended in April 2018 at the Florida International University. This work was advised by Professor Leonardo Bobadilla from the same institution.

The revolution of autonomous vehicles has led to the development of robots with abundant sensors, actuators with many degrees of freedom, high-performance computing capabilities, and high-speed communication devices. These robots use a large volume of information from sensors to solve diverse problems. However, this usually leads to a significant modeling burden as well as excessive cost and computational requirements. Furthermore, in some scenarios, sophisticated sensors may not work precisely, the real-time processing power of a robot may be inadequate, the communication among robots may be impeded by natural or adversarial conditions, or the actuation control in a robot may be insubstantial. In these cases, we have to rely on simple robots with limited sensing and actuation, minimal onboard processing, moderate communication, and insufficient memory capacity. This reality motivates us to model simple robots such as bouncing and underactuated robots making use of the dynamical system techniques. In this dissertation, we focus on four broad themes to solve problems in resource-constrained scenarios: 1) Combinatorial filters for bouncing robot localization; 2) Bouncing robot navigation and coverage; 3) Stochastic multi-robot area patrolling; and 4) Deployment and planning of underactuated aquatic robots.

The striking motivation for the approaches in this dissertation is the global analysis of simple robotics systems. This global analysis of robotics systems leads us to use a dynamical system technique. The dynamical system we use here is the *cell-to-cell mapping* technique (originally introduced by Hsu) (Hsu, 1980, 2013). In the cell-to-cell mapping, the state space is divided into small cells, where each cell is considered a state entity. In our approaches, we utilize two cell-to-cell mapping techniques which are the *simple cell-to-cell mapping* (SCM) and the *generalized cell-to-cell mapping* (GCM). In the SCM, each cell has only one image cell. In the GCM, each cell has several image cells. The GCM is a generalization of the SCM. The modeling of the deterministic behavior of robots leads to the application of the SCM. The formulation of the nondeterministic behavior of robots in terms of the GCM leads to a finite Markov chain. These dynamical system techniques provide the attractors (limit cycles) and domains of attraction from the system behavior which allowed us to develop the filters, controllers, and algorithms for the solutions to localization, navigation, coverage, planning, patrolling, and deployment problems. In the following, the summaries of four contributions in broad research themes are explained.

Combinatorial Filters for Bouncing Robot Localization Mobile robot localization is the problem of determining a robot's configuration (position and orientation) in its environment, and it is typically a prerequisite to solving other robotic problems. The motivation of our work is to use a robot with limited linear and angular sensing as a basis for investigating the intrinsic limits of the localization problem. In the first contribution of the dissertation, we focus on a setup that considers a known polygonal environment with obstacles and a robot equipped only with a clock and contact (or bump) sensors called a *bouncing robot*. We consider that the bouncing robot has access to a map of its environment, but is initially unaware of its position and orientation within that environment. This bouncing robot is modeled in a predictable way: the robot moves in a straight line and then bounces from the environment's boundaries by rotating in place counterclockwise through a bouncing angle. The problem of global robot localization is how the robot deduces its con-

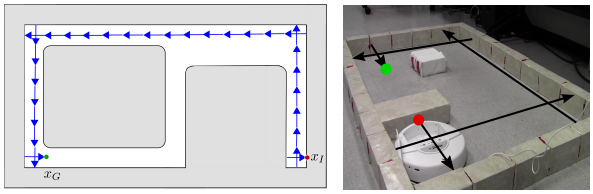
figuration following its modeled behavior. Can this bouncing robot be globally localized without even knowing its initial configuration? Using our setup, we present a global localization method for a bouncing robot (Alam, Bobadilla, & Shell, 2018). Our method finds the limit cycles and their transient trajectories from a known environment using the SCM and generates I-state graphs. We then use these I-state graphs to synthesize filters to solve the localization problem. Our localization filters take less computation time and memory compared to traditional Bayesian filter-based localization approaches (Thrun, Fox, Burgard, & Dellaert, 2001; Fox, 2003; Leonard & Durrant-Whyte, 1991). Figure 1(a) shows two limit cycles generated from our simulation and a corresponding physical experiment for the bouncing robot localization is illustrated in Figure 1(b).



(a) (b)

Figure 1: Localization from limit cycles.

Bouncing Robot Navigation and Coverage In the second contribution, we use the same bouncing robot model to investigate both the navigation and coverage problems once the localization problem is solved. The problem of navigation is finding a path for a robot between an initial configuration and a goal configuration. The coverage problem of the environment is visiting all locations of interest using one or more robots. How could the simple behavior of the bouncing robot be useful in solving the common robotic problems, such as navigation and coverage, with limited linear and angular sensing? In multi-robot settings, will many such bouncing robots be useful as well to solve the coverage problem? Our proposed solution (Alam, Bobadilla, & Shell, 2017) in this contribution has the following steps: 1) A directed graph is constructed from the environment geometry based on the GCM from the simple bouncing policies. 2) The shortest path on the graph, for navigation, is generated between either one given pair of initial and goal configurations or all possible pairs of initial and goal configurations. 3) The optimal distribution of bouncing policies is computed so that the actual coverage distribution is as close as possible to the target coverage distribution. A simulation result and a physical experiment of the navigation path between a given pair of initial and goal configurations in the environment are shown in Figure 2. In this contribution, we also create a sampling-based joint trajectory of multiple bouncing robots incrementally to cover the given environment starting from an initial configuration instead of going over all the states in the high-dimensional state space.

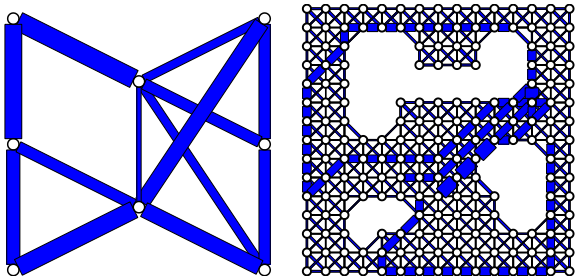


(a) (b)

Figure 2: Bouncing-based navigation.

Stochastic Multi-Robot Area Patrolling In the third contribution, we investigate the problem of area patrolling in an adversarial situation in which a number of robots as patrollers visit a group of locations of interest in an environment to detect the intrusion of an adversary. In a communication-constrained and adversarial environment, it is a challenging problem for multiple robots to patrol the whole environment by sensing with their limited ability to see. In the multi-robot patrolling problem, what will be an efficient method for robots to patrol an area under the adversarial scenario? How can we remove the need for synchronization and coordination among the patrolling robots? How can the robots with limited visibility be used to patrol an adversarial and communication-constrained environment? Deterministic patrolling strategies could also be learned by an adversary observing them over time. Therefore, we alternately use randomized patrolling strategies based on Markov chains for several reasons: 1) These will make it harder for an adversary to successfully complete an attack and evade its detection due to the unpredictability of the strategies. 2) A randomized motion can be

easily implemented in a mobile robot, since its communication, sensing, and computation requirements are minimal. 3) Efficient algorithms can calculate Markov chains with the desired properties (Ghosh, Boyd, & Saberi, 2008). In this contribution, we propose distributed patrolling strategies for guarding a set of locations in an environment under adversarial attacks (Alam, Edwards, Bobadilla, & Shell, 2015; Alam, 2016) and present a method of finding patrolling policies for multiple patrollers that guard any polygonal environment using limited visibility regions and nondeterministic paths (Alam, Rahman, Bobadilla, & Rapp, 2017). The randomized patrolling policies for patrollers in different environments represented as graphs are illustrated in Figure 3, where the width and color saturation of edges are proportional to the optimal edge weight value or the probability of that edge being chosen by a patroller.



(a) (b)

Figure 3: Randomized patrolling policies.

Deployment and Navigation of Underactuated Aquatic Robots In the final contribution, we are interested in tackling the problem of deploying multiple underactuated aquatic robots called *drifters* so that their desired long-term trajectories can gather aquatic data visiting all locations on the surface of a marine environment. We also tackle the problems of path planning and finding navigation policy for the drifter. The drifters drift passively with ambient ocean currents. Vertical actuation (buoyancy) enables them to alter their depth and achieve controllability by the use of different current layers in the ocean. How can we model the behavior of the drifter in a marine environment? In addition, the study of a marine environment is a challenging task because of the spatiotemporal variations of ocean phenomena and the disturbances caused by ocean currents. As such, we must collect data from a marine environment over long periods of time to better assess and understand a marine environment. The uncertainty of the drifter motion due to the disruption of ocean currents and winds needs to be taken into account in our motion model of the drifter. In this contribution, we present a data-driven, deployment and navigation approach for the drifters. We extract the generalized flow pattern within a given region from ocean model predictions, develop a Markov chain-based motion model using the GCM, and analyze the long-term water flow behavior. Based on this long-term behavior of the water flow, we find a minimum number of deployment locations for the drifters in the marine environment (Alam, Reis, Bobadilla, & Smith, 2018). A generated vector field from the Regional Ocean Modeling System (ROMS) (Shchepetkin & McWilliams, 2005) predicted oceanic current data in the Southern California Bight (SCB) region, California, USA, is shown in Figure 4(a). We found attractors and their transient groups or the domains of attraction of the environment as the long-term behavior of the water flow. The initial deployment locations of the drifters based on this long-term behavior are illustrated in Figure 4(b). All possible reachable locations from an initial deployment location of the drifter are determined as its planned, long-term drifter trajectory. An optimal navigation policy is developed to demonstrate the best possible action from any location to a goal location in the environment.

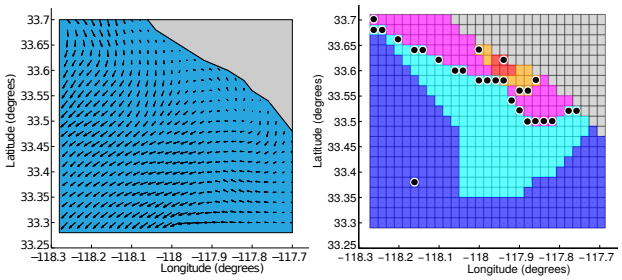
Acknowledgments

The author would like to acknowledge the financial support of a Florida International University Graduate School Dissertation Year Fellowship. This work was also supported in part by Army Research Office Grant 67736CSII.

References

Alam, T. (2016, December). Decentralized and nondeterministic multi-robot area patrolling in adversarial environments. *International Journal of Computer Applications*, *156*(2), 1–8.

Alam, T., Bobadilla, L., & Shell, D. A. (2017). Minimalist robot navigation and coverage using a dynamical system approach.



(a) (b)

Figure 4: Data-driven deployment: (a) The vector field generated from ROMS current prediction data; (b) Attractors (blue and red regions) and associated transient groups (the cyan region for the blue attractor, the orange region for the red attractor, and the magenta region for both attractors); The initial deployment locations (filled in black circles) for regions of these attractors and transient groups.

In *Proceedings of the IEEE International Conference on Robotic Computing* (pp. 249–256).

Alam, T., Bobadilla, L., & Shell, D. A. (2018, January). Space-efficient filters for mobile robot localization from discrete limit cycles. *IEEE Robotics and Automation Letters*, *3*(1), 257–264.

Alam, T., Edwards, M., Bobadilla, L., & Shell, D. (2015). Distributed multi-robot area patrolling in adversarial environments. In *the International Workshop on Robotic Sensor Networks.*

Alam, T., Rahman, M. M., Bobadilla, L., & Rapp, B. (2017). Multi-vehicle patrolling with limited visibility and communication constraints. In *Proceedings of the IEEE Conference on Military Communications* (pp. 465–470).

Alam, T., Reis, G. M., Bobadilla, L., & Smith, R. N. (2018). A data-driven deployment approach for persistent monitoring in aquatic environments. In *Proceedings of the IEEE International Conference on Robotic Computing* (pp. 147–154).

Fox, D. (2003). Adapting the sample size in particle filters through KLD-sampling. *International Journal of Robotics Research*, *22*(12), 985–1003.

Ghosh, A., Boyd, S., & Saberi, A. (2008). Minimizing effective resistance of a graph. *SIAM Review*, *50*(1), 37–66.

Hsu, C. S. (1980). A theory of cell-to-cell mapping dynamical systems. *Journal of Applied Mechanics*, *47*(4), 931–939.

Hsu, C. S. (2013). *Cell-to-cell mapping: a method of global analysis for nonlinear systems* (Vol. 64). Springer Science & Business Media.

Leonard, J. J., & Durrant-Whyte, H. F. (1991). Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, *7*(3), 376–382.

Shchepetkin, A. F., & McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, *9*(4), 347–404.

Thrun, S., Fox, D., Burgard, W., & Dellaert, F. (2001). Robust Monte Carlo localization for mobile robots. *Artificial Intelligence*, *128*(1-2), 99–141.

Tauhidul Alam received his B.Sc. in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Bangladesh, and his Ph.D. in Computer Science from Florida International University. His research interests span the areas of Robotics, Motion Planning, and Cyber-Physical Systems. In particular, his research aims to devise solutions to the fundamental robotic tasks of localization, navigation, coverage, and patrolling of mobile robots with limited sensing, actuation, computation, and communication.