CIS 700-002: Topics in Safe Autonomy, Spring 2019

Reading List (under construction)

Simulators and Verification Tools

Driving Simulators

- CARLA: An Open Urban Driving Simulator. Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltu\ n. 1st Conference on Robot Learning (CoRL 2017). (www.carla.org)
- AirSim. Microsoft. (https://www.microsoft.com/en-us/research/blog/autonomous-car-research/) (https://github.com/Microsoft/AirSim)
- ROS + Gazebo
- Unity's SimViz Solution Template
- PreScan Simulation Platform

Verification Tools

- "dReach: δ-Reachability Analysis for Hybrid Systems." Kong, Soonho, Sicun Gao, Wei Chen and Edmund M. Clarke. TACAS (2015).
- "Flow*: An Analyzer for Non-Linear Hybrid Systems." Chen X., Abraham E., Sankaranarayanan S. In: Sharygina N., Veith H. (eds) Computer Aided Verification. CAV 2013. Lecture Notes in Computer Science, vol 8044. Springer, Berlin, Heidelberg.

Papers

Autonomous Systems

- "Autonomous Systems--An Architectural Characterization." Joseph Sifakis. 2018. (https://arxiv.org/abs/1811.10277)
- "Formal Methods for Semi-Autonomous Driving." Sanjit A. Seshia, Dorsa Sadigh, S. Shankar Sastry . Proceedings of the Design Automation Conference (DAC), June 2015 (https://people.eecs.berkeley.edu/~dsadigh/Papers/seshia-dac2015.pdf)

Anomaly Detection

- "Outlier Analysis" (2nd ed.) 2016. Aggarwal, Charu C. Springer Publishing Company, Incorporated.
- "Anomaly detection: A survey." Varun Chandola, Arindam Banerjee, Vipin Kumar. CSUR '09 (https://dl.acm.org/citation.cfm?id=1541882)
- "Precision and Recall for Time Series." Nesime Tatbul, Tae Jun Lee, Stan Zdonik, Mejbah Alam, Justin Gottschlich. NeurIPS 2018 (https://arxiv.org/abs/1803.03639/)
- "Security of Cyber-Physical Systems in the Presence of Transient Sensor Faults." Junkil Park, Radoslav Ivanov, James Weimer, Miroslav Pajic, Sang Hyuk Son, and Insup Lee. ACM Transactions on Cyber-Physical Systems: 1(3), 2017.
- "Greenhouse: A Zero-Positive Machine Learning System for Time-Series Anomaly Detection." Tae Jun Lee, Justin Gottschlich, Nesime Tatbul, Eric Metcalf, Stan Zdonik. SysML '18 (https://arxiv.org

/abs/1801.03168)
- "A Generalized Zero-Positive Learning System to Detect Software Performance Anomalies." Mejbah Alam, Justin Gottschlich, Abdullah Muzahid. AutoPerf (https://arxiv.org/abs/1709.07536)
- "Production-run software failure diagnosis via Adaptive Communication Tracking." Mohammad Mejbah ul Alam, Abdullah Muzahid . ISCA '16 (https://dl.acm.org/citation.cfm?id=3001175)
- "Evaluating Real-Time Anomaly Detection Algorithms - The Numenta Anomaly Benchmark." Alexander Lavin, Subutai Ahmad. IEEE ICMLA, 2015. (https://arxiv.org/abs/1510.03336)
- "Demystifying Numenta Anomaly Benchmark." Nidhi Singh, Craig Olinsky. IJCNN, 2017. (https://ieeexplore.ieee.org/abstract/document/7966038)
- "Paranom: A Parallel Anomaly Dataset Generator." (DATSA) Justin Gottschlich. (https://arxiv.org/pdf/1801.03164.pdf)

Data Set Shift

- Moreno-Torres, Jose G., et al. "A unifying view on dataset shift in classification." Pattern Recognition 45.1 (2012): 521-530. [pdf]
- Sugiyama, Masashi, Neil D. Lawrence, and Anton Schwaighofer. Dataset shift in machine learning. The MIT Press, 2017. [pdf] **Particularly I.1, II.3, and III.8**
- Klinkenberg, Ralf, and Thorsten Joachims. "Detecting Concept Drift with Support Vector Machines." ICML. 2000. [pdf]
- Raza, Haider, Girijesh Prasad, and Yuhua Li. "EWMA model based shift-detection methods for detecting covariate shifts in non-stationary environments." Pattern Recognition 48.3 (2015): 659-669. [pdf]

Medical Applications

- "Not to Cry Wolf: Distantly Supervised Multitask Learning in Critical Care." Patrick Schwab, Emanuela Keller, Carl Muroi, David J. Mack, Christian Strassle, Walter Karlen. ICML. 2018. [pdf]
- "Reducing Pulse Oximetry False Alarms Without Missing Life-Threatening Events." Hung Nguyen, Sooyong Jang, Radoslav Ivanov, Christopher P. Bonafide, James Weimer, Insup Lee. Smart Health. 2018. [pdf]

Verification of Neural Networks and Closed-Loop Systems

- Verification of Hybrid Systems
  - "Taylor Model Flowpipe Construction for Non-linear Hybrid Systems". Xin Chen, Erika Abraham, Sriram Sankaranarayanan. (https://www.cs.colorado.edu/~xich8622/papers/rtss12.pdf)
  - "dReach/dReal". Soonho Kong, Sicun Gao, Wei Chen, and Edmund Clarke. (https://scungao.github.io/papers/dreach_tool.pdf)
- Verification of neural networks
  - "Reluplex: An Efficient SMT Solver for Verifying Deep Neural Networks*". Guy Katz, Clark Barrett, David Dill, Kyle Julian and Mykel Kochenderfer. May, 2017. (https://arxiv.org/pdf/1702.01135.pdf)
  - "Output Range Analysis for Deep Neural Networks". Souradeep Dutta,Susmit Jha, Sriram Sanakaranarayanan, Ashish Tiwari. Sep, 2017. (https://arxiv.org/pdf/1709.09130.pdf)
- Verification using Abstract Interpretation
  - "AI2: Safety and Robustness Certification of Neural Networks with Abstract Interpretation". Timon Gehr, Matthew Mirman, Dana Drachsler-Cohen, Petar Tsankov, Swarat Chaudhuri , Martin Vechev. IEEE S&P 2018. (https://files.sri.inf.ethz.ch/website/papers/sp2018.pdf)
  - "Differentiable Abstract Interpretation for Provably Robust Neural Networks". Matthew Mirman, Timon Gehr, Martin Vechev. ICML 2018. (https://files.sri.inf.ethz.ch/website/papers/icml18-diffai.pdf)
  - "Fast and Effective Robustness Certification". Gagandeep Singh, Timon Gehr, Matthew

Mirman, Markus Puschel, Martin Vechev. NIPS 2018. (https://papers.nips.cc/paper/8278-fast-and-effective-robustness-certification.pdf)
    - "Robustness Certification with Refinement". Anonymous authors. Accepted for ICLR19. (https://openreview.net/pdf?id=HJgeEh09KQ)
- Verification/Falsification of closed-loop systems with NN components
    - "Verisig: Verisig: verifying safety properties of hybrid systems with neural network controllers". Radoslav Ivanov, James Weimer, RAjeev Alur, George J. Pappas, Insup Lee. (https://arxiv.org/pdf/1811.01828.pdf)
    - "Formal Verification of Neural Network Controlled Autonomous Systems." X. Sun, H. Khedr, Yasser Shoukry, Oct 2018. (https://arxiv.org/pdf/1810.13072)
    - "Compositional Falsification of Cyber-Physical Systems with Machine Learning Components." Tommaso Dreossi, Alexandre Donze, Sanjit A. Seshia. Dec 2018. (https://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-165.pdf)
    - "Simulation-based Adversarial Test Generation for Autonomous Vehicles with Machine Learning Components." Cumhur Erkan Tuncali, Georgios Fainekos, Hisahiro Ito, James Kapinski. Jan 2019. (http://www-bcf.usc.edu/~jdeshmuk/teaching/cs699-fm-for-cps/Papers/C4.pdf)
    - "Reasoning about Safety of Learning-Enabled Components in Autonomous Cyber-physical Systems." Cumhur Erkan Tuncali, James Kapinski, Hisahiro Ito, Jyotirmoy V. Deshmukh. Apr 2018. (https://arxiv.org/pdf/1804.03973.pdf)

Runtime Verification

- Assumptions
- Reachability

Testing

- "DeepXplore: Automated Whitebox Testing of Deep Learning Systems." Kexin Pei, Yinzhi Cao, Junfeng Yang, Suman Jana, SOSP 2017.
- "Simulation-based Adversarial Test Generation for Autonomous Vehicles with Machine Learning Components." Cumhur Erkan Tuncali, Georgios Fainekos, Hisahiro Ito, James Kapinski. IEEE Intelligent Vehicles Symposium (IV) 2018.
- "APEX: Autonomous Vehicle Plan Verification and Execution." M. O'Kelly, H. Abbas, S. Gao, S. Shiraishi, S. Kato, R. Mangharam, 2016. (http://repository.upenn.edu/mlab_papers/84)

Self-Driving Vehicles

- "On a Formal Model of Safe and Scalable Self-driving Cars." Shai Shalev-Shwartz, Shaked Shammah, Amnon Shashua. Mobileye, 2017. (https://arxiv.org/pdf/1708.06374.pdf)
- "Self-Driving Vehicle Verification Towards a Benchmark." Nima Roohi, Ramneet Kaur, James Weimer, Oleg Sokolsky, Insup Lee. June 2018. (https://arxiv.org/abs/1806.08810)
- "Interaction-aware occupancy prediction of road vehicles." M. Koschi and M. Althoff. 20th IEEE International Conference on Intelligent Transportation Systems, pages 1885-1892, 2017. (http://www.i6.in.tum.de/Main/Publications/Koschi2017b.pdf)

Medical Applications: Smart Alarms & Closed-loop Verification

- Not to Cry Wolf: Distantly Supervised Multitask Learning in Critical Care Patrick Schwab, Emanuela Keller, Carl Muroi, David J. Mack, Christian StrÂ¨assle, Walter Karlen. PMLR 2018.
- Reducing Pulse Oximetry False Alarms Without Missing Life-Threatening Events. Hung Nguyen, Sooyong Jang, Radoslav Ivanov, Christopher P. Bonafide, James Weimer, and Insup Lee. In Proceedings of IEEE/ACM Conference on Connected Health: Applications, Systems and

Engineering Technologies (CHASE 2018), Washington, D.C., USA, September 2018.
- An Intraoperative Glucose Control Benchmark for Formal Verification. Sanjian Chen, Matthew O'Kelly, James Weimer, Oleg Sokolsky, and Insup Lee. In 5th IFAC conference on Analysis and Design of Hybrid Systems (ADHS 2015), Atlanta, GA, USA, October 2015.
- Data-driven Adaptive Safety Monitoring using Virtual Subjects in Medical Cyber-Physical Systems: A Glucose Control Case Study. Sanjian Chen, Oleg Sokolsky, James Weimer, and Insup Lee. In Journal of Computer Science and Engineering, Volume 10, Num.3, pp.75-84, September 2016 (Open Access).

Interpretable Machine Learning

- "The Mythos of Model Interpretability." Zachary C. Lipton. 2016. (https://arxiv.org/abs/1606.03490)
- "Towards A Rigorous Science of Interpretable Machine Learning." Finale Doshi-Velez, Been Kim. 2017. (https://arxiv.org/abs/1702.08608)

Human-in-the-Loop / Behavior Modeling

- "A Data-Driven Behavior Modeling and Analysis Framework for Diabetic Patients on Insulin Pumps." Sanjian Chen, Lu Feng, Michael R. Rickels, Amy Peleckis, Oleg Sokolsky, and Insup Lee. IEEE International Conference on Healthcare Informatics 2015 (ICHI 2015), October 2015.
- "Data-Driven Probablistic Modeling and Verification of Human Driver Behavior." Dorsa Sadigh, Katherine Driggs Campbell, Alberto Alessandro Angelo Puggelli, Wenchao Li, Victor Shia, Ruzena Bajcsy, Alberto L. Sangiovanni-Vincentelli, S. Shankar Sastry, Sanjkit Seshia. Formal Verification and Modeling in Human-Machine Systems, AAAI Spring Symposium, March 2014. (https://people.eecs.berkeley.edu/~dsadigh/Papers/sadigh-fvmhs2014.pdf)
- "Synthesis for Human-in-the-Loop Control Systems." Wenchao Li, Dorsa Sadigh, S. Shankar Sastry, Sanjit Seshia. 20th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS), April 2014. (https://people.eecs.berkeley.edu/~dsadigh/Papers/li-tacas2014.pdf)

PLC Security

- Simon Duque Anton, Lia Ahrens, Daniel Fraunholz, and Hans D. Schotten. Time is of the Essence:Machine Learning-based Intrusion Detection inIndustrial Time Series Data. ICDMW, Nov 2018. (https://arxiv.org/pdf/1809.07500.pdf)
- Huan Yang, Liang Cheng, Mooi Choo Chuah: Detecting Payload Attacks on Programmable Logic Controllers (PLCs). CNS 2018: 1-9
- Luis Garcia, Ferdinand Brasser, Mehmet Hazar, Osama Mohammed, Ahmad-Reza Sadeghi, Saman Zonouz, Hey, My Malware Knows Physics! Attacking PLCs with Physical Model Aware Rootkit Network, Network and Distributed System Security Symposium (NDSS), 2017.
- Luis Garcia, Saman Zonouz, Dong Wei, Leandro Pfleger de Aguiar, Detecting PLC Control Corruption via On-Device Runtime Verification, IEEE Resilience Week 2016.

Infrastructure/Platform Support

- Middleware
- Edge Computing

Assurance Techniques

- "Combining Software Evidence - Arguments and Assurance." R. Weaver, G. Despotou, T. Kelly, J. McDermid. ACM REBSE'05 (https://www-users.cs.york.ac.uk/tpk/REBSE05.pdf)

Last updated on 2/6/19 by Taylor Carpenter.