Artificial Intelligence in Automated Vehicle Systems: Risk Mitigation Strategies in Emerging Technology

July 30, 2019

The field of artificial intelligence (AI) has expanded rapidly in the past decade. AI has already transformed how we shop, how we unlock our phones, how we experience social media, and how involved we are when driving. Like any other emerging technology, AI is also experiencing growing pains, from chat bots and facial recognition algorithms with racial and gender biases to avionics complications. Issues with automated vehicle (AV) systems may follow, with potentially serious consequences for users and manufacturers.

Existing standards for vehicles, robotics, and software development focus on traditional programming techniques that make system behaviors the direct consequences of specific choices made by designers or programmers, sometimes traceable to a specific line of code. In an AV system using AI, this may no longer be the case. Instead, the system is designed to learn appropriate behavior from training scenarios. Although programmers may direct development by designating desirable and undesirable behavior, ultimately many system behaviors may not be directly attributable to a specific programmer decision or line of code. This deviation from traditional programming renders many best practices insufficient for validating safe performance and justifying performance decisions in the wild. It also demands critical examination of training scenarios to ensure vehicles behave as desired even in new territories.

AV systems may exhibit undesirable behavior when faced with conditions not fully explored in training scenarios, such as encountering inclement weather or pedestrians using different cultural norms to signal crossing (for example, making eye contact with the "driver" versus using hand signals to indicate whether they will cross or wait for the vehicle to proceed). In the absence of commonly accepted standards and regulations, original equipment manufacturers (OEMs) and AI developers must determine for themselves how

best to engineer systems, write software, train and validate systems, and mitigate risks for AV providers and other roadway users.

How to Ensure Best Practices

Efforts are underway to help AV system developers ensure best practices and mitigate risks. However, regulations and policies will not develop at the same pace as the technology. OEMs and AI developers must take responsibility for establishing best practices to address the unique challenges presented by non-traditional programming and systems development requirements applied to AVs. Developing common industry standards in this domain efficiently will require collaboration among leaders in the AV community, rather than each provider developing its own internal standards.

The collision avoidance metrics partnership (CAMP) provides a template for addressing issues of self-regulation in emerging technologies. From the mid-1990s into the 2000s, OEMs worked in partnership to solve numerous safety issues, including developing out-of-position-occupant air bag safety, side impact collision requirements, front impact collision compatibility requirements, Vehicle-to-All communications protocols (V2X), and initial pre-competitive research into advanced driver assist technologies. A similar collaborative effort

can be applied to AVs, but the Silicon Valley AV providers must be involved with traditional OEMs to address challenges specific to these new technologies.

Although the release of new standards and guidance may be slow, there are several strategies OEMs and software developers can adopt in the meantime. For example, OEMs and developers can combine best practices from industries outside of the vehicle and software development spaces, including cybersecurity, the financial services sector, and software quality standards. By drawing on cross-disciplinary expertise, AV developers will be better poised to avoid issues of bias in their training data (pulling on methods based in classical statistics frequently applied in the banking and financial industries), address consumer privacy concerns, and handle systems integration issues.

Liability

As with human drivers, the question of liability when AVs crash will be important for economic damage coverage; in AVs, should fault attach to the vehicle, the conditions that led to the conflict must be assessed, analyzed, and integrated into improved controls to avoid recurrence. The first step in determining liability will be determining whether "the AI" or "the AV" is at fault. This requires immaculate and detailed records of the incident from an event data recorder (EDR). After a crash, investigators may compare data from the EDRs of multiple vehicles with other observations to fully understand the circumstances that led to the event. The prevailing operating conditions that preceded a collision (including system readiness, data, and data processing) and whether an appropriate notice or control action had been issued will need to be registered in an EDR and available to enable safety researchers to attribute causation and, where necessary, promote engineering of corrective actions and remedy.[1] Where an AV is at fault, public confidence in AVs may be adversely affected.

Event records will be reviewed critically from multiple parties looking to determine whether the system had previously encountered the scenario leading to the crash event and whether it should have and/or could have been trained to handle the scenario differently. It is the responsibility of AV developers to curate training and system validation data to ensure and demonstrate that the vehicle can safely operate in the prevailing environment. In addition, AV developers must ensure that the system gracefully disengages when exposed to unfamiliar scenarios and must revert to a fail-safe condition that does not initiate safety challenges for other roadway users.

Identifying relevant dangerous conditions outside the training set (sometimes called corner or edge cases) is an ongoing challenge that will continue as AVs are rolled out in new environments, under new conditions,

and with increasingly less human interaction. If the training dataset is not diverse enough AVs may not be able to function under new environmental conditions, recognize actions of other roadway users, or respond appropriately to pedestrian behaviors and customs. Although AV developers and OEMs are motivated to be first to market, collaboration and cross-disciplinary expertise are the best way to expose vulnerabilities in their systems and curate sufficiently diverse training sets. AV developers and OEMs must also stay apprised of and participate in developing new standards and regulations.

Recommendations

Take the time to design and implement best practices now, before litigation proceeds from crashes in the wild.

Share edge cases among members of a pre-competitive cooperative research project or consortium, using CAMP as a model. Consider forming industry consortia to develop common standards for function and data recording.

Practice traditional risk mitigation strategies, such as DFMEAs (Design Failure Mode and Effects Analyses) using multi-disciplinary teams including software developers, cybersecurity experts, human factors experts, and statisticians. Critically examine diversity of training sets.

Roll out slowly, anticipate unexpected behaviors in new environments and register learnings to be 1) tested in simulation for all possible relevant use cases and 2) updated from the AV provider to the fleet operating in the wild.

Use an objective third party to review code, training data, vehicle performance, and simulations for validation testing. Evaluate, share, and use machine-learning test tools to scrutinize training data and AI and AV performance.

Stay apprised of current and upcoming policies and regulations. Be part of the process and propose rules to meet the need for motor vehicle safety. Consider formation of collaborative partnerships to perform and report upon precompetitive research of common interest, develop commonly accepted industry standards for software, security, event data logging, identification and sharing rare events (street sweepers path variations), establish common signaling for interactions with vulnerable road users (cyclists, pedestrians) and to satisfy other needs as become identified. Exponent has expertise in vehicles, systems integration, robotics, human interaction, traditional programming, and AI. We can provide objective, third-party review, custom vehicle test plans, and guidance on existing and emerging standards (including Safety of the Intended Function or SOTIF, ISO 26262, and NHTSA guidelines). Exponent can work with multiple clients to: form partnerships, identify common research needs and interests, select research topics, timely deliver research results, protect client Intellectual Work Product and ensure confidentiality among competitors.

---

[1] Lange et al., "Data Requirements for Post-Crash Analyses Of Collisions Involving Collision Avoidance Technology Equipped, Automated, And Connected Vehicles," Paper 17-0338, ESV 2017.

Exponent has the expertise and experience to:

- Write vehicle, systems, and vehicle level requirements related to general and specific use-case applications. Simulate and test those requirements in development, training, and validation phases.

- Consult upon and/or engineer AV systems including optics; vision data processing and conversion; sensor performance and sensor fusion; low-voltage power systems; high-voltage battery and battery controls systems; electrical system and PCB layout review; and systems engineering, development and verification/validation planning.

- Cyber Security.

- HMI/user experience, acceleration/deceleration profiles, jerk, lateral Gs, optimization for occupant comfort, and maximized passenger/user experiences.

- ISO 26262, Functional Safety, and ISO/PAS 21448, Safety of the Intended Functionality

- Electrical engineering standards and applicability to compute and control systems.

- Third-party code review and software development process review.

- Autonomous Vehicle Event Data recording content.

- Process FMEA (PFMEA) and Design Review Based on Failure Mode (DRBFM)

Exponent can perform research and generate intellectual work product within multiple organizational frameworks including but not limited to individual client institutions, pre-competitive informal research consortia, a consortium as an individual legal entity, and multiple project sponsors (in which multiple institutions with common interests jointly sponsor and fund research and share the resultant intellectual work product equally).

Any emerging technology introduces new risks. Automated vehicle (AV) systems present new challenges but practicing traditional risk mitigation strategies (such as DFMEAs and software verification and validation) using multi-disciplinary teams and interrogating AVs with machine-learning-based tools and emerging common industry standards can mitigate potentially serious consequences for users and help drive AVs through early application stages.

Robert C. Lange
Vehicle Engineering
Senior Fellow, Principal
Menlo Park

(248) 324-9107
rlange@exponent.com

Joel C. Wilson, Ph.D., P.E.
Vehicle Engineering
Managing Engineer
Menlo Park

(650) 688-7165
jwilson@exponent.com

Elizabeth Groves, Ph.D.
Electrical Engineering & Computer Science
Managing Scientist
Menlo Park

(650) 688-7147
egroves@exponent.com

Exponent Office Locations
Atlanta, Austin, Boston Area (Maynard, Natick), Chicago Area (Downtown Chicago, Warrenville), Denver, Detroit, Houston, Miami, New York, Philadelphia, Phoenix, Northern California Area (Menlo Park, Oakland, Sacramento), Seattle, Southern California Area (Los Angeles, Orange County, Pasadena), Washington DC Area (District of Columbia, Maryland, Virginia)

International Offices:
Basel, Switzerland; Derby, Harrogate and London, UK; Düsseldorf, Germany; Shanghai and Hong Kong, China

E^xponent®