Ethical Concerns with Autonomous Vehicles

or

“Why your car will be programmed to kill you!”

Thomas P. Pannett, Esq., P.E., MBA, CPPO
Administrator, ODOT Office of Contract Sales
Division of Construction Management
1980 W. Broad St.
Columbus, Ohio
614-466-2739
Tom.Pannett@dot.ohio.gov
OUTLINE

• Definitions
• History and Evolution of Artificial Intelligence
• A.I. Ethics
• Evolution of Autonomous Vehicles
  • 5 Stages
  • Where they are now? (Tesla vs Google)
• Ethical Issues of Autonomous Vehicles
• 15 Points Checklist (#11)
• Asimov’s Laws
• Ethical Dilemma Survey
• Why Cars Programmed to Kill You

“Ethics is knowing the difference between what you have a right to do and what is right to do.” (Potter Stewart)
• AI or “Artificial Intelligence” is the theory and development of computer systems able to perform human cognitive-related tasks.

• AV (automated or autonomous) According to the USDOT: An automated vehicle system is a combination of hardware and software (both remote and on-board) that performs a driving function, with or without a human “actively monitoring” the driving environment.

• CV (connected vehicle) is a car that shares internet access internally an externally.

• Engineering Ethics is the study of applied ethics and system of moral principals used in the practice of engineering.

• Ethical Dilemma is a decision-making problem between two moral imperatives; neither of which is completely preferable.
Lies, Damn Lies & Statistics

- Approximately 40,000 drivers/passengers die each year in auto accidents
  - That is 1.18 fatalities for every 100 million miles traveled
  - Americans drive nearly 3.2 trillion miles each year

- About 6,000 pedestrians die each year in automobile accidents

- Only a few have died in autonomous vehicle accidents to date

- 90%+ of accidents can be attributed to human error

- If avoided, could save hundreds of billions in time and cost and loss of life

- BUT ..... It would take 15 years to replace the current non-driverless fleet with autonomous vehicles if we started today.

The fact that man knows right from wrong proves his intellectual superiority to the other creatures; but the fact that he can do wrong proves his moral inferiority to any creatures that cannot.”  — Mark Twain
Evolution of AI

- AI was conceptualized in Science Fiction Novels in the early 1900s.
- The first mathematical models showing that AI was possible were developed in the 1950s.
- From 1957 to 1974, AI flourished. Computers stored more data becoming faster & cheaper.
- In the 1980s, AI was enhanced via algorithmic toolkits and started to develop self-learning techniques.
- In 1997, IBM’s Deep Blue was able to beat Gary Kasparov, master chess champion.
- NOW: big data and massive computing abilities allow computers to learn through brute force.
- To be truly “intelligent,” AI needs to continue to evolve and is doing so through the comprehension of language.
- The ultimate goal of AI is AGI (artificial general intelligence); meaning, not just smart and fast, but the ability to perform any intellectual task a human could do.
- Once AGI is achieved, ASI (artificial super intelligence) is next; when a super fast thinking machine develops cognitive abilities and multiplies them and its understanding of things so quickly (because it is a computer) that it will have super-intelligence evolving exponentially.

“Any tool can be used for good or bad. It’s really the ethics of the artist using it.”  
(John Knoll)
History of AI

“Any tool can be used for good or bad. It’s really the ethics of the artist using it.”

(John Knoll)

ARTIFICIAL INTELLIGENCE TIMELINE

1938-1946: Golden Age of Science Fiction
1950: Can Machines Think? - Alan Turing
1956: Dartmouth Summer Research Project on Artificial Intelligence
1955: Logic Theorists, the first AI program, is invented
1963: DARPA funds AI at MIT
1965: Moore’s Law
1968: “By the year 2001 we will have machines with intelligence that matched or exceeded human’s” - Arthur Clarke and Steve Kubrik
1970: “From 3-8 years we will have a machine with the general intelligence of a human being” - M. Minsky
1980: Edward Feigenbaum introduces expert systems
1982: Japan’s Fifth Generation Computer Project
1986: Navlab, the first autonomous car, is built by Carnegie Mellon
1997: Deep Blue defeats Gary Kasparov in chess
1997: First publicly available speech recognition software developed by Dragon Systems
Evolution of AI Ethics

• Current AI algorithms with human-equivalent or superior performance are characterized by a deliberately programmed competence only in a single, restricted domain ... think Deep Blue’s Chess Match / programming

• Generality is missing (i.e., AGI).

• As AGI is constructed, it must be programmed to act in many domains, with many consequences, including problems the engineers never explicitly envisioned.

• Thus AGI MUST be explicitly programmed to imbue good behavior in such terms as “X such that the consequence of X is not harmful to humans.”

• And because AGI will conceive of alternatives not yet dreamed, Ethical cognition itself must be taken as a subject matter of engineering. We must “engineer ethics.”

Modern bureaucrats often take refuge in established procedures that distribute responsibility so widely that no one person can be identified to blame for the catastrophes that result. Howard, 1994
• Many big-name companies, like Google, already have ethics boards in place to monitor the development and deployment of their AI. But, the details of who sits on Google’s ethics board and what it actually does remains scarce.

• Recently, Facebook, Google, and Amazon launched a consortium that develop solutions to AI safety and privacy issues.

• “OpenAI” is an organization dedicated to developing and promoting open-source AI for the good of all (openai.com).

• One promising branch of AI development is using Bayesian inference for ethical decision-making.
  • The theorem’s developed focus on a self-modifying and evolutionary program that improve a decision (i.e., make it more ethical or safe-human centric) as more evidence becomes available about the problem.


“To do injustice is the greatest of all evils.” (Plato)
How important is the AI Ethics discussion?

• “A.I. is a bigger concern than terrorism or climate change ....”
  • Professor Jim Al-Khalili, head of the British Science Association

• “If Russian cyber hackers were able to meddle with the 2016 US elections, then what is stopping cyber terrorists from hacking into any future AI controlled power grids, transport systems, banks of military installations,” he continued. “Our government has a responsibility to protect society from potential threats and risks.”

• “The danger of AI is much greater than the danger of nuclear warheads by a lot.” Elon Musk

BUT SURELY, NO ONE WOULD PURPOSLY CREATE MALIGNANT A.I. .... Rrriiiggghth!

DATELINE JUNE 2018: Researchers at MIT created a “psychopathic” A.I. it dubbed Norman. During an ink blot test, “a ‘normal’ A.I. perceived ‘a black and white photo of a small bird,’ Norman thought it looked more like a man who ‘gets pulled into a dough machine.’
• A.I. is here and A.G.I. is coming;

• A.I. can be created for good or bad, and already is being created for both;

• The community of users (governments and engineers especially) need to be engaged in the discussion of A.I. Ethics;
  • A common determinate for an ethical direction must be made before AGI

• A.I. = probably the most important part of A.V./C.V.
Evolution of Autonomous Vehicles

Currently, 5 “Levels” of Automation
(Adopted by SAE International)

• **Level 1**: Automated system on the vehicle occasionally assists the human driver conduct some parts of the driving task;
  • Most if not all vehicles currently made have at least Level 1 automation – think “anti-lock” breaks

• **Level 2**: Automated system on the vehicle can actually conduct some parts of the driving task, while the human continues to monitor the driving environment and performs the rest of the driving task;
  • Think “adaptive cruise control” and “active lane assist.” Tesla currently produces at this level.

• **Level 3**: Sometimes called “Conditional Automation” … system conducts some parts of the driving task and monitors the driving environment in some instances, but the human driver must be ready to take back control when the automated system requests;
  • This was the level of the Uber vehicle from 2018 that was in the pedestrian accident in Tempe, AZ.

• **Level 4**: “Hight Automation” … System can conduct the driving task and monitor the driving environment, and the human need not take back control, but the automated system can operate only in certain environments and under certain conditions; and

• **Level 5**: “Full Automation”… System performs all driving tasks under all conditions.

“More often there’s a compromise between ethics and expediency.” (Peter Singer)
Society of Automotive Engineers (SAE)

0. No Automation
   Zero autonomy; the driver performs all driving tasks.

1. Driver Assistance
   Vehicle is controlled by the driver, but some driving assist features may be included in the vehicle design.

2. Partial Automation
   Vehicle has combined automated functions, like acceleration and steering, but the driver must remain engaged with the driving task and monitor the environment at all times.

3. Conditional Automation
   Driver is a necessity, but is not required to monitor the environment. The driver must be ready to take control of the vehicle at all times with notice.

4. High Automation
   The vehicle is capable of performing all driving functions under certain conditions. The driver may have the option to control the vehicle.

5. Full Automation
   The vehicle is capable of performing all driving functions under all conditions. The driver may have the option to control the vehicle.
Where they are now? (Tesla vs Google)

- “Waymo” (Goolge) recently logged its 5 millionths mile traveled.
  - Performing tests with hundreds and soon thousands of vehicles in the Phoenix area
  - Actually have Level 4 in use but on a mapped area
- Uber recently reached 2 million miles
- Tesla’s numbers aren’t counted because it is not considered autonomous

The Ethical Person always seeks excellence in one’s self or one’s station. (Aristotle)
"Be as you wish to seem."  (Socrates)
15 Points Safety Assessment Checklist

https://www.transportation.gov/sites/dot.gov/files/docs/AV%20policy%20guidance%20PDF.pdf

National Highway Traffic and Safety Administration (part of the U.S. DOT)

• Safety Assessment Letter to NHTSA with 15 Point Checklist for manufacturers and others to voluntarily follow and report upon.
• Suggesting they report to the Office of the Chief Counsel for each HAV system, outlining how they are meeting this Guidance at the time they intend their product to be ready for use (testing or deployment) on public roads.
• The Safety Assessment would cover the following areas:
  1) Data Recording and Sharing;
  2) Privacy;
  3) System Safety;
  4) Vehicle Cybersecurity;
  5) Human Machine Interface;
  6) Crashworthiness;
  7) Consumer Education and Training;
  8) Registration and Certification;
  9) Post-Crash Behavior;
  10) Federal, State and Local Laws;
  11) Ethical Considerations;
  12) Operational Design Domain;
  13) Object and Event Detection and Response;
  14) Fall Back (Minimal Risk Condition);
  15) Validation Methods
A. Decisions made by the AV computer “driver” will have ethical dimensions or implications.
   i. Do I turn left into the building or ram the vehicle ahead of me?

B. Different outcomes for different road users may flow from the same real-world circumstances.
   i. Will GMC have me ram the building where Ford will direct me to hit the vehicle ahead of me?
   ii. Will it matter what jurisdiction I’m in (California vs. Ohio vs. France vs. China?)

(a) Computer will be programmed with decision rules and learning;
(b) The programming of an HAV may establish an implicit or inherent decision rule with significant ethical consequences; and
(c) Manufacturers and other entities, working cooperatively with regulators and other stakeholders (e.g., drivers, passengers and vulnerable road users), should address these situations to ensure that such ethical judgments and decisions are made consciously and intentionally.
   (1) Who are the “other entities” and when will we get our chance to be a stakeholder?

C. Algorithms for resolving conflict situations should be developed transparently using input from Federal and State regulators, drivers, passengers and vulnerable road users, and taking into account the consequences of an HAV’s actions on others.
AV Ethics: Still more questions that answers!

• Can an automobile be programmed to follow traffic laws?
  • Yes, b/c mostly prescriptive … speed, lane limitations, access limitations, signalization, signage interpretation, etc. … See, ORC 4511 et seq.

• Can an automobile be programmed to avoid an accident?
  • Eventually … but …
    • Ultimately, in my opinion, it will have to be all or “nothing” to really reduce human error.
    • In other words, to achieve the goals of AV/CV (less congestion, safer roads, etc.), we will have to adopt full autonomy or at least Level 4. Thus, the human element will be almost completely removed and we will have to rely on the computer to decide.

• Will an automobile in China be programmed differently than an automobile in the U.S.?
  • Laws are different … certainly societal ethics are different

• Can an ethical frameworks derived from humans be formed into the algorithms that control cars?
Will AV/CVs follow the “Laws?”

ASIMOV’S THREE LAWS OF ROBOTICS

1. A ROBOT MAY NOT INJURE A HUMAN BEING OR, THROUGH INACTION, ALLOW A HUMAN BEING TO COME TO HARM.

2. A ROBOT MUST OBEY ORDERS GIVEN TO IT BY HUMAN BEINGS, EXCEPT WHERE SUCH ORDERS WOULD CONFLICT WITH THE FIRST LAW.

3. A ROBOT MUST PROTECT ITS OWN EXISTENCE AS LONG AS SUCH PROTECTION DOES NOT CONFLICT WITH THE FIRST OR SECOND LAW.
0th Law: “A robot may not harm humanity, or, by inaction, allow humanity to come to harm.”

VIKI and the “4th Law”

VIKI from “I Robot” ....

“To protect Humanity, some humans must be sacrificed. To ensure your freedom, some freedoms must be surrendered. We robots will ensure mankind's continued existence. You are so like children. We must save you from yourselves.”

“Do the right thing because it is right.” (Immanuel Kant)
1. An automated vehicle **should not** collide with a pedestrian or cyclist.

2. An automated vehicle **should not** collide with another vehicle except where avoiding such a collision which would conflict with the First Law.

3. An automated vehicle **should not** collide with any other object in the environment, except where avoiding such a collision would conflict with the First or Second Law.

4. An automated vehicle **must** obey traffic laws, except where obeying such laws would conflict with the first three laws.
Lets talk about the “Trolley Problem”
Ted Talk, What moral decisions should driverless cars make?

Research by Iyad Rahwan

https://www.ted.com/talks/iyad_rahanwwhat_moral_decisions_should_driverless_cars_make

www.moralmachine.mit.edu
Video 1

Use societal trends based on underpinnings of an ethical mores to direct the research and decision making of the moral machine.

https://youtu.be/XPN1kHKcQ-Q

What would Kant do?
How Does Ethical Dilemma Survey Work?

Q(a): Would you sacrifice yourself for a group or would you hit the small group?

Q(b): Would you swerve to kill someone or let the vehicle hit the person in the cross-walk?

Q(c): Would you swerve to sacrifice yourself for the Large Group?
“Excellence is never an accident. It is always the result of high intention, sincere effort, and intelligent execution; it represents the wise choice of many alternatives – choice, not chance, determines your destiny.” (Aristotle)
Video 2

Do people taking the moral machine survey want the regulation necessary to have a car make decisions for them and would they buy one?

https://youtu.be/1YX5yx2fTw
Perception vs Reality

• 73% of people are afraid to ride in a fully-autonomous vehicle

• Only 40% of people believe autonomous vehicles are safer than regular cars
  • Went down after Uber accident when it was 50%

• Over 80% people believe they are an above average safe driver

• Up to 60% of drivers would not choose an AV if offered the option

• Most drivers will sacrifice themselves to save others, but almost all drivers will sacrifice strangers before family

“Good people do not need laws to tell them to act responsibly, while bad people will find a way around the laws.” (Plato)
The German Resolution

• According to researches in Germany:
  
  • In the event of **unavoidable accident** situations, **any distinction** based on personal features (age, gender, physical or mental constitution) is **strictly prohibited**. It is also **prohibited to offset** victims against one another. General programming to **reduce the number** of personal injuries **may** be justifiable. Those parties involved in the generation of mobility risks must **not sacrifice non-involved parties**.

  • In the case of **automated** and connected driving systems, the **accountability** that was previously the sole preserve of the individual shifts from the motorist **to the manufacturers** and **operators** of the technological systems **and to the bodies** responsible for taking infrastructure, policy and legal decisions. Statutory **liability** regimes and their fleshing out in the everyday decisions taken by the courts **must** sufficiently **reflect** this transition.

• Translated:
  
  • The scenarios in which an AV would have to make the life or death split decision are so complex and unpredictable that there is no way for ethical decisions to be standardized in programming code or regulation.
  
  • In other words, they have gone around the “Gordian Knot”:
    … no trolley problem … not allowed to program … avoidance is primary … if not, the situation will be evaluated on a case by case basis … is this helpful?
  
  • Is this what Kant would do? Yes, follow the **moral law** and not dissimilar to Asimov.

  

  “Crime and the criminal confront us with the perplexity of radical evil; but only the hypocrite is really rotten to the core.”  **Hannah Arendt**
Our “enemies” lack ethics

“We are what we repeatedly do. Excellence, then, is not an act, but a habit.” (Aristotle)
Cars Will Be Programmed to Kill

Hyperbole? Yes, but it will happen.

But Why?

- $$$ ... Follow the money; if it ultimately is less expensive (insurance, cost of death, road delays, regulations, etc.), it will happen.

- Because they will be programmed to avoid accidents at all cost, and it will be cheaper and safer to slam you into a wall than to hit another vehicle.

  - In other words, the vehicle would seek to minimize damage in a global sense in the event of a dilemma situation, thereby reducing the societal impacts of accidents.

- But, will you be able to override (and thus shift the ethical decision & all liability back to you)?

  - That is a decision we as potential policy makers in the transportation industry need to help make.

Pascal’s wager: Choose to live as if God exists, for if he does not, you are out very little, but if he does exist, your bad choices will cost you eternity.
THANK YOU

The “Ethical Engineer” makes each decision using his/her highest mental capacity, diligence and skill while continually keeping the best interest of the public in mind.

-- me

Thomas P. Pannett, Esq., P.E., MBA, CPPO
Administrator, ODOT Office of Contract Sales
Division of Construction Management
1980 W. Broad St.
Columbus, Ohio
614-466-2739
Tom.Pannett@dot.ohio.gov