# Autonomous Systems

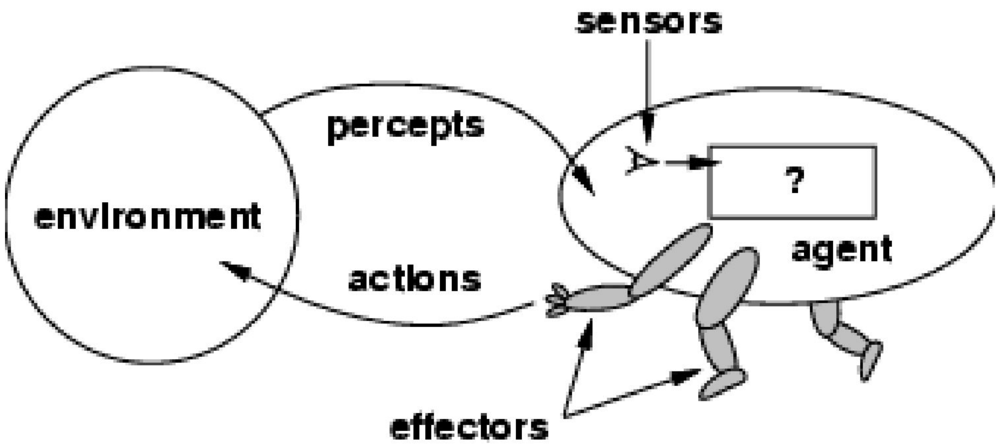
Adane Letta, PhD

January 2025

# Outline

- Autonomous navigation and decision-making
- Case studies: Autonomous vehicles, drones, and industrial robots
- Ethical considerations in autonomous robots

# Autonomous Navigation and Decision-Making




Autonomous navigation and decision-making are critical aspects of robotics and artificial intelligence (AI), enabling systems to navigate environments and make decisions without human intervention.

# Autonomous Navigation and Decision-Making

- What is Autonomous Navigation?

Autonomous navigation is the ability of a system (e.g., a robot, drone, or vehicle) to move through an environment without external guidance while avoiding obstacles, following paths, and achieving predefined goals.

- Key Components
  - Perception: Gathering data from the environment using sensors such as cameras, LiDAR, radar, and GPS.
  - Localization: Determining the system's position within the environment (e.g., Simultaneous Localization and Mapping, or SLAM).
  - Mapping: Building a map of the environment using sensor data.
  - Path Planning: Calculating an optimal and collision-free route.
  - Control Systems: Executing navigation commands to follow the planned path.

# Autonomous Navigation and Decision-Making

## Challenges in Autonomous Navigation

1. Sensor Reliability: Errors due to bad weather, low light.
2. Ethical Considerations: Decision-making in life-and-death situations.
3. Regulatory Hurdles: Legal approvals for autonomous systems.
4. High Computational Requirements: Processing vast amounts of data in real-time

# Autonomous Navigation and Decision-Making

## Decision-Making in Autonomous Systems

- Refers to the process by which an autonomous system selects optimal actions to achieve its goals, such as reaching a destination, avoiding obstacles, or following rules, in dynamic and potentially unpredictable environments.

- Decision-making involves selecting optimal actions based on the system's goals, constraints, and sensory inputs.

# Autonomous Navigation and Decision-Making

- Decision-Making in Autonomous Systems
  - Key Components:
    - Perception: collecting and interpreting sensory data (e.g., cameras, LiDAR, radar, GPS).
    - Reasoning: analyzing sensory inputs to determine the best course of action.
    - Planning: Short-term (local) and long-term (global) route planning.
    - Control: Executing the planned actions through control systems.

# Algorithms for Decision-Making

Dynamic programming and probabilistic techniques play a significant role in autonomous navigation and decision-making, particularly in handling complex, dynamic, and uncertain environments.

# Dynamic Programming (DP) Principle

- DP involves solving complex problems by breaking them down into simpler sub-problems.
- Useful for optimization and sequential decision-making.
- There may not be a “standard form” of DP problems, instead is an approach to problem solving and algorithm design

# Dynamic Programming (DP) Principle

- Start with a small portion of the problem and find optimal solution for this smaller problem

- Gradually enlarge the problem – finding the current optimal solution from the previous one

  … until original problem is solved in its entirety

- This general philosophy is the essence of the DP principle
  - The details are implemented in many different ways in different specialised scenarios

# Characteristics of DP Problems

- DP problems all share certain features:

1. The problem can be divided into stages, with a policy decision required at each stage

2. Each stage has several states associated with it

3. The effect of the policy decision at each stage is to transform the current state into a state associated with the next stage (could be according to a probability distribution, as we'll see next).

Policy decision refers to the choice made at each stage or step of the problem to optimize the outcome.
- Define how to transition from one state to another while aiming for the best result based on the problem's objective (e.g., minimizing cost, maximizing profit, etc.).

# Characteristics of DP Problems

5. Given the current state, an optimal policy for the remaining stages is **independent** of the policy adopted in previous stages
6. The solution procedure begins by finding the optimal policy for each state of the last stage.
7. Recursive relationship identifies optimal policy for each state at stage $n$, given optimal policy for each state at stage $n+1$:

$$f_n^*(s) = \min_{x_n}\{c_{sx_n} + f_{n+1}^*(x_n)\}$$

8. Using this recursive relationship, the solution procedure moves backward stage by stage – until finding optimal policy from initial stage

# DP Algorithms

- Bellman Equation: Defines the optimal value of a decision as the immediate reward plus the future reward from subsequent decisions.
  - Used in path planning to calculate the shortest or most cost-efficient path.

- Value Iteration: Iteratively calculates the value of each state to converge on an optimal policy.
  - Used in Markov Decision Processes (MDPs) to make optimal decisions under uncertainty

  - Common in robot navigation where dynamic changes (e.g., moving obstacles) must be considered

# DP Algorithms

- Policy Iteration: Alternates between policy evaluation and improvement to find the best decision policy.
    - Suitable for systems like self-driving cars, where both safety and efficiency are critical.
- Dynamic Time Warping (DTW): Matches time-dependent sequences.
    - DTW aligns system movement with desired paths, even when there are timing mismatches.

# Markov Decision Model

- Probabilistic techniques are vital for handling uncertainty in real-world environments, such as unpredictable obstacles, sensor noise, and dynamic changes.

- Consider the following application: machine maintenance
- A factory has a machine that deteriorates rapidly in quality and output and is inspected periodically, e.g., daily
- Inspection declares the machine to be in four possible states:
  - 0: Good as new
  - 1: Operable, minor deterioration
  - 2: Operable, major deterioration
  - 3: Inoperable
- Let $X_t$ denote this observed state:
  - Evolves according to some "law of motion", it is a stochastic *process*
  - Furthermore, assume it is a finite state Markov chain

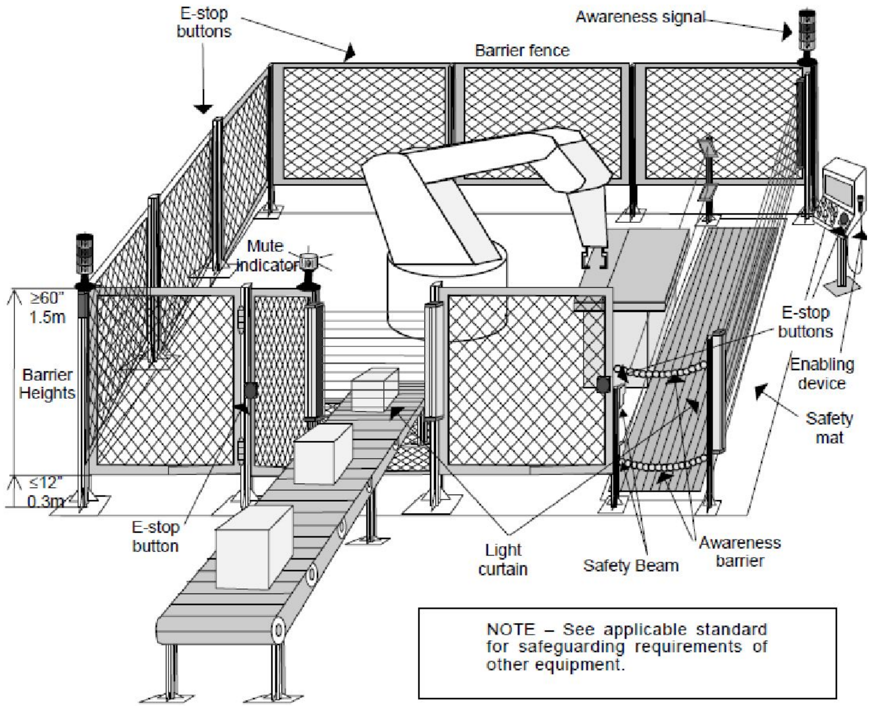
# Markov Decision Model

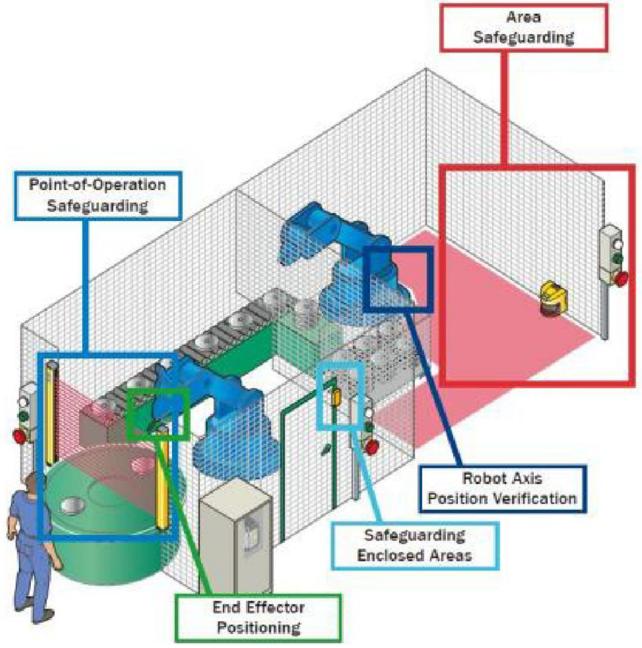
- Markov Decision Processes (MDPs): Models environments with states, actions, transition probabilities, and rewards.
  - Decision-making in uncertain environments (e.g., route planning under traffic conditions).
- Bayesian Inference: Updates the probability of a hypothesis as more evidence is obtained.
  - Estimating the likelihood of collisions or system failures.

# Markov Decision Model

- Monte Carlo Methods: Uses random sampling to estimate numerical results.
    - Monte Carlo Localization (MCL): Helps robots estimate their location by sampling possible positions and refining them based on sensor data.
    - Monte Carlo Tree Search (MCTS): Used in decision-making for real-time, high-dimensional problems (e.g., gaming strategies or multi-robot collaboration)
- Hidden Markov Models (HMMs): Models systems where states are not directly observable but can be inferred through observable outputs.
    - Recognizing patterns in sensor data (e.g., pedestrian movement prediction).

# Notions of Robot Safety



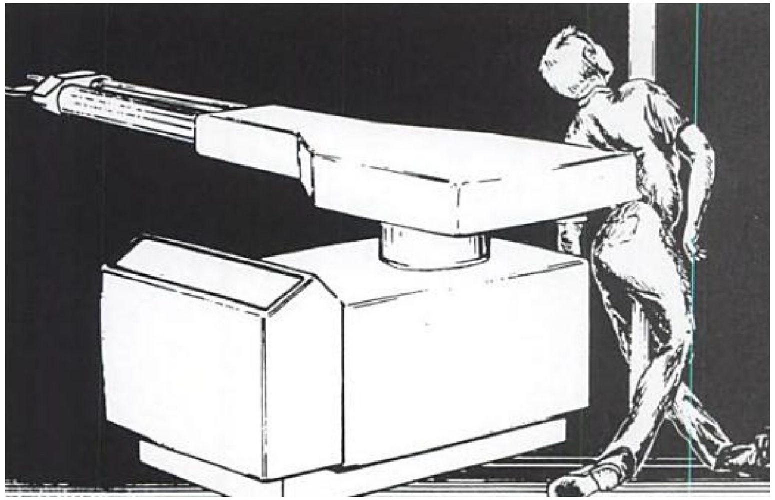

- Robots, depending on the task, may generate paint mist, welding fumes, plastic fumes, etc.
- The robot, on occasion is used in environments or tasks too dangerous for workers, and as such creates hazards not specific to the robot but specific to the task.

# Examples of Accidents

- **Example 1: First fatal robot-related accident in the U.S**.

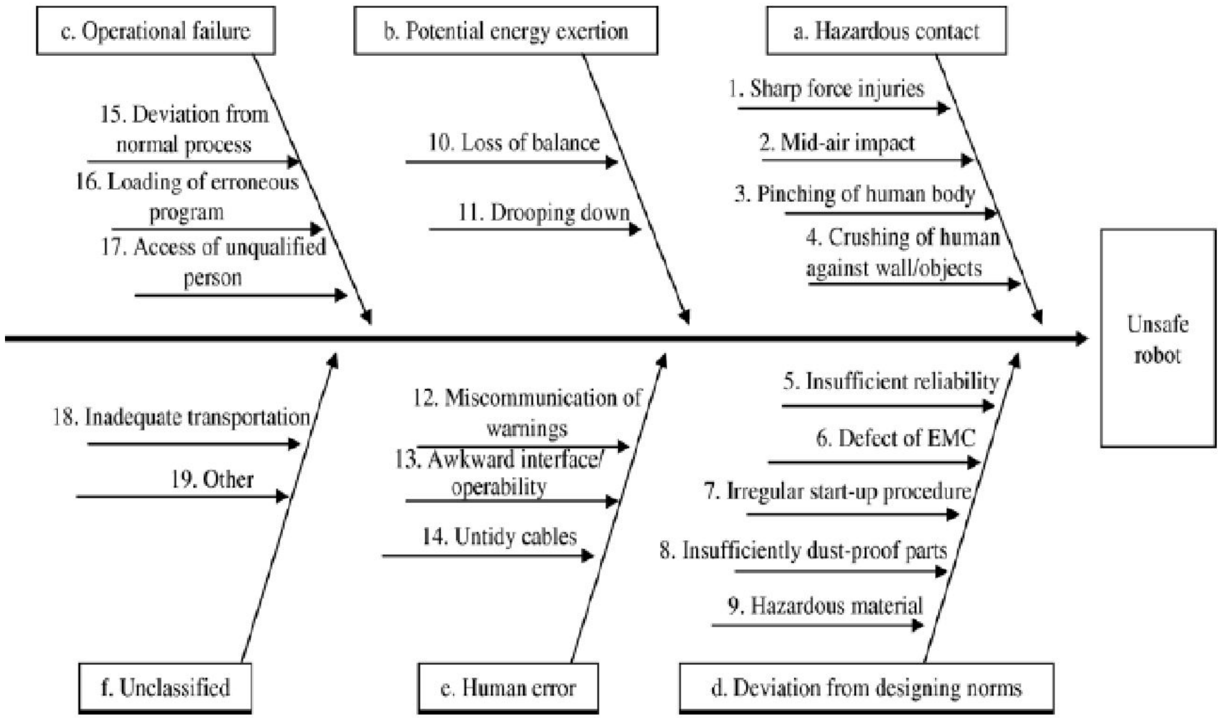
- On July 21, 1984, a die cast operator was working with an automated die cast system utilizing a Unimate Robot, which was programmed to extract the casting from the die-cast machine, dip it into a quench tank and insert it into an automatic trim press.

- A neighboring employee discovered the victim pinned between the right rear of the robot and a safety pole in a slumped but upright position. The victim died five days later in the hospital.



# Characterizing an Unsafe Robot

- Human Interaction
- Control Errors
- Unauthorized Access
- Mechanical Failures
- Environmental Sources
- Power Systems
- Improper Installation



# Safety Cases

- The purpose of a safety case:

A safety case should communicate a clear, comprehensive and defensible argument that a system is acceptably safe to operate in a particular context

- Safety cases are already adopted in many industries, including defence, aerospace, railways and automotive sectors.

- Based on such practice, we can extract a few key attributes of what makes a good and useful safety case

# Aspects of a Safety Case

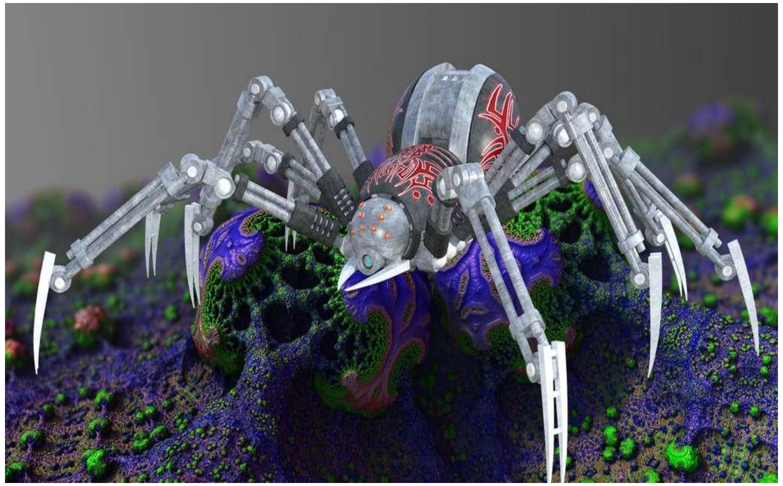
- Argument: the case exists to communicate an argument, to demonstrate how someone can reasonably conclude that a system is acceptably safe from the available evidence
- Clear: it is a device for communicating ideas and information to a third party, e.g., regulator
- System: this could be anything from a network of pipes to a software module with parameters or operating procedures
- Acceptably: In most applications, “absolute” safety is impossible. So, the case is argument to say how the system is safe enough (as per some notion of tolerable risk)
- Context: most systems are only safe within a context of use, which should be defined in the safety case

# Safety Case as a Physical Artifact

- Comprehensive and structured set of documentation
- To ensure safety can be demonstrated with reference to:
  - Arrangements and organisation, including emergency arrangements
  - Safety analyses
  - Compliance with standards and best practice
  - Acceptance tests
  - Audits and inspections
  - Feedback

# Case studies: Autonomous vehicles, drones, and industrial robots

- Animat:- refers to an autonomous robot
- To exhibit common sense a robot must be able to:
  - Manipulate models of its world,
  - Reason by analogy,
  - Carry out useful line of reasoning automatically, and
  - Develop enough of a "self" to sensible modify its own behavior.



- Artificial life (AL) is a different approach to artificial intelligence…. based on the following line of reasoning.
  - The idea is that we can understand adaptive behavior best if we focus on the way a behaving individual and its environment interact.

# Case studies:….

- The behaviors of living beings are not the simple sum of the behaviors of all their component parts but are the results of the interaction of those parts.

- These interactions resulted in an emergent behavior

- *“Intelligence can only be determined by the total behavior of the system and how that behavior appears in relation to the environment.”* (Brooks, 1991)

# Case studies:….

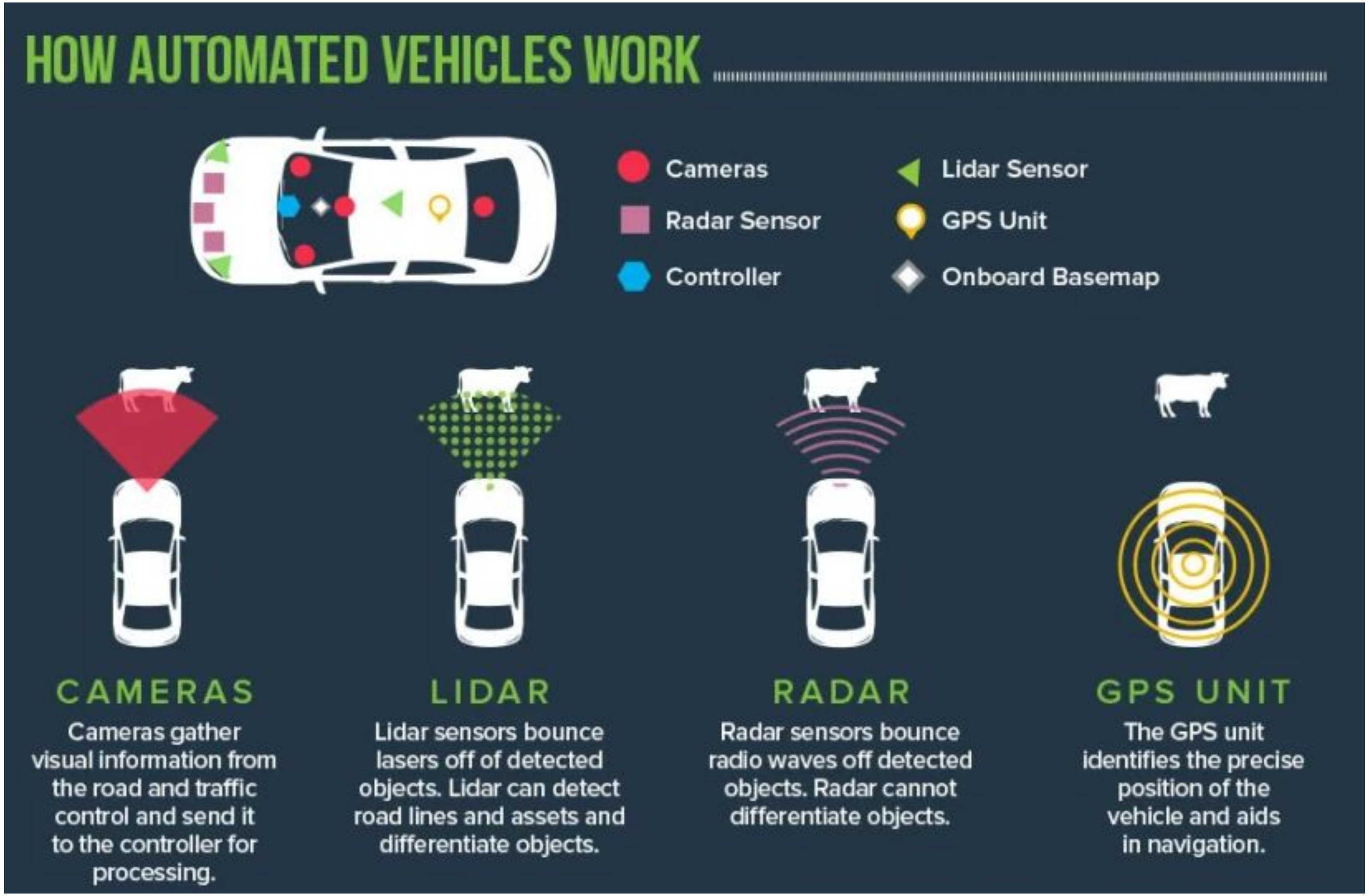
- Motivation for placing animats in real situation derives from a consideration of what real animals do:
  - Animals confronted with a complex environment and multiple tasks to satisfy
  - Animals behave continuously and react to stimuli without significant delay for planning
    - Sensor and motor systems in animals are connected by multiple parallel connections that vary in speed and complexity of response.
  - Animals typically possess some basic behaviors from birth, and then develop more complex abilities, including many high-level cognitive abilities, including motor skills

# Case studies:….

- There are over 3 million robots worldwide
- Autonomous vehicles
  - In 1939, General Motors presented the first-ever autonomous car at the New York World’s Fair. Was
    - Powered by electricity
    - Steered by radio-controlled electromagnetic fields
  - Why?
    - Currently there are about 1.3 million traffic deaths worldwide in one year
    - Traffic accident happened mostly due to human error
    - “Autonomous vehicles can easily avoid all these”

There are questions on how autonomous vehicles should behave, and how responsibility and risk should be distributed in the complicated system the vehicles operates in.

# Case studies:….



# Case studies:….

- There are six levels of autonomy:
  - Level 0… cars have no autonomy---- 100% driven by humans
  - Level 1… cars have driving control share between humans and AI (e.g. parking assistant)
  - Level 2… cars have automated system that execute acceleration, steering and braking … human driver monitors the environment
  - Level 3… resemble to level 2 except that drivers can turn their attention from driving tasks
  - Level 4 … cars are controlled by automated systems when the driving conditions are especially difficult ….. E.g. sever weather
  - Level 5 … cars have total autonomy

# Case studies:….

- The great majority, so-called, autonomous vehicles …. Produced by Tesla, General Motors, Volvo and Mercedes are only level 2
  - Human driver is constantly monitoring and supervising their functioning.

- Enormous complexities involved in producing safe truly autonomous vehicles ---- may not happen soon.

- Are autonomous vehicles safer than those driven by humans?.. Current state.

# Case studies:….

- The RAND Corporation (2016), estimated autonomous vehicles would have to drive approximately 11 billion miles to draw reliable conclusions compared to human drivers.

- Worrying straws in the mind
    - The first death involving an autonomous vehicle, Tesla Model S electric car, happened on the 7 May 2016.

- How are most autonomous vehicles driven?

# Case studies:….

- Autonomous vehicles are driven differently from human-driven cars…. slower
  - Many accidents happened in autonomous vehicles are rear-end crashes due to their unusual slowness …. % of rear-end crashes is more than twice as high compared with humans
  - Human drivers make inaccurate prediction about movement of autonomous cars
  - Autonomous vehicle takes 0.83 seconds on average following disengagement for the human driver to take over control
    - A car driven at 50 mph covers approximately 60 feet in 0.83 seconds ….. Lead to fatality

# Case studies:….

- How should autonomous vehicles be programmed?
  - Two main ways:
    - Modular system: associated with perception (track moving object, detect traffic signals) and decision making (route planning, behavior selection--- traffic light handling, intersection handling
    - Porammed via deep learning or reinforcement learning –
      - Deep learning leads to flexibility in driving behavior --- used in real-world driving situations …. Disadvantage not poor generalizability to novel situation

# Case studies:….

- Koopman et. al (2019) mentioned that autonomous vehicles need to be programmed to respond appropriately to novel situations. …. many autonomous vehicles fall short of fulfilling the criterion.

- “In the future autonomous cars will become safer than human-driven ones”

# Case studies:….

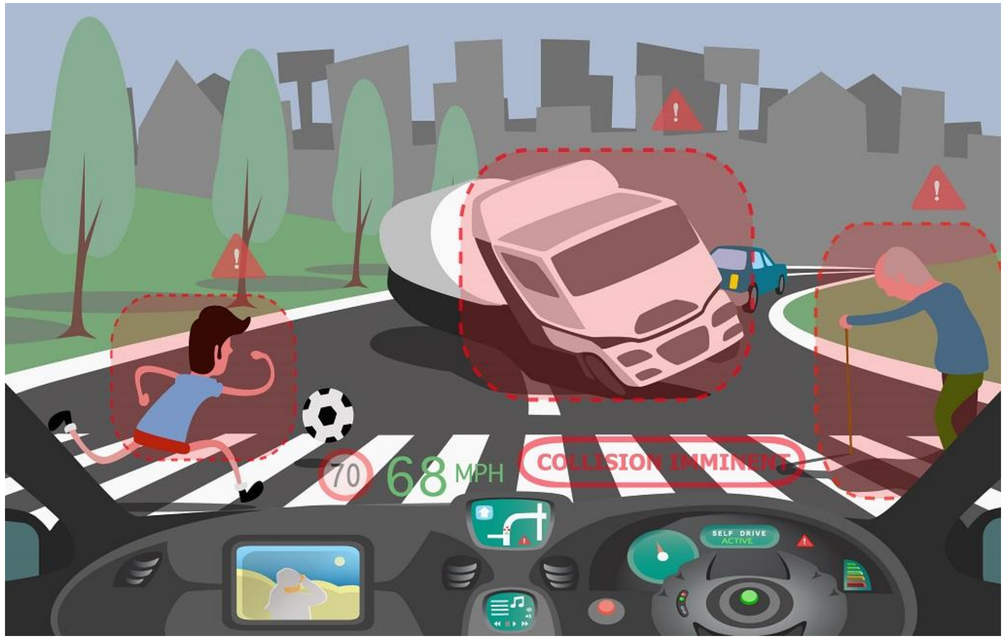
- Some technologies, like nuclear power, cars, or plastics, have caused ethical and political discussion and significant policy efforts to control the trajectory these technologies, usually only once some damage is done.

- Once we have understood a technology in its context, we need to shape our societal response, including regulation and law.

- Plus the more fundamental fear that they may end the era of human control on Earth

# Robot and Morality

- Ethical issues
  - There is a debate among politicians, AI experts, psychologists, and philosophers on autonomous vehicles…. Should be programmed ethically.
  - The "trolley problems" used as a theoretical tool to investigate ethical intuitions and theories
  - Jeremy Bentham (1748-1832)…. Utilitarianism…. The greatest good for the greatest number



# Robot and Morality

- Warfare-autonomous weapons …. killer robots  and other lethal AI systems
    - Robots and other autonomous weapons systems will play a central role in the future wars.
        - In 2008, there were between 4,000 – 6,000 ground robots in Iraq…. Used to detonate roadside explosive devices
        - Since then MQ-9 Reaper unmanned aerial vehicle has been developed …. As a killer robots
    - Role of AI in warfare
        1. Allowing machines to act without human supervision
        2. Processing and interpreting large amounts of data ‘
        3. Aiding command and control of war

The main arguments against (lethal) autonomous weapon systems are that they support extrajudicial killings, take responsibility away from humans, and make wars or killings more likely

# Robot and Morality

- Using robots in warfare has several potential advantages
  - Very efficient … need few human controller to monitor of the use of thousands of autonomous weapons

  - Robots are more ethical than humans

  - Why?

# Robot and Morality

- Using robots in warfare has several potential advantages
  - Because they never become angry or seek revenge

  - Robots can be programmed to conform to the human-made laws governing warfare
- Potential disadvantages
  - Their use could trigger a global arms race
  - Autonomous weapon systems are becoming increasingly complex…..probability of accidents occurring increases as system complexity increases
  - Human cannot calculate precisely what will happen whey they are deployed
    - Explainable AI….. providing transparent accounts for their proposed decisions to ensure they are consistent with human decision making …… hard to achieve

Explainable AI does not mean that we expect an AI to “explain its reasoning”— would require far more serious moral autonomy than we currently attribute to AI systems.

# Robot and Morality

- Potential disadvantages
  - Hacking is potentially a huge problem …. Hackers could potentially design a virus that would reverse the programming so the drone killed only civilians

- Moral issues
  - “ a missile strike on a terrorist compound even though it would risk a child’s life or to cancel the strike to protect the child but thereby risk a terrorist attack”
  - Human controller is responsible to the command structure

The lack of transparency of many killer robots and the relative ease with which their programming can be hacked pose substantial moral issues

# Robot and Morality

- Moral issues
    - Another question is whether using autonomous weapons in war would make wars worse, or make wars less bad.

# Robot and Morality

- Robots as a “new social species”
  - “How robots become members of a cyber-physical social network that is cohabited with sapient humans”
  - “The challenge is to explicate this process whereby sociality, which is inherently a human-to-human trait, is forced to negotiate itself in a cyber-physical world shared with humanoids (human-like robots), animaloids (animal-like robots) and a myriad of socially-interactive intelligent machines.”
  - Attempts had been made to theorizing human–robot sociality “through the dynamics of interpretation, signification and attribution”

Lamola, M.J. On the Robosphere: A Philosophical Explication of the Socio-technical Status of Social Robots. *Int J of Soc Robotics* **14**, 1199–1209 (2022). https://doi.org/10.1007/s12369-021-00840-3

# AI and Morality

- Does AI have moral agency?
  - There has been a steady increase in the number of people killed by robots or other AI-systems
    - Autonomous vehicles
    - Robots in a factory
  - Why AI is perceived as morally responsible?
  - Who has moral responsibility for these and other AI-based tragedies?
    - Is it the AI system?
    - Is it the designers of the AI system? Or
    - Both?

# Ethics of AI and Robotics

- EU policy document suggests "trustworthy AI" should be lawful, ethical, and technically robust, and then spells this out as seven requirements:
  - Human oversight,
  - Technical robustness,
  - Privacy and data governance,
  - Transparency,
  - Fairness,
  - Well-being, and
  - Accountability
- On the other hand, Perhaps a "code of ethics" for AI engineers, analogous to the codes of ethics for medical doctors

Stanford Encyclopedia of Philosophy (2020). Ethics of Artificial Intelligence and Robotics

# Ethics of AI and Robotics

- Manipulation of Behaviour
  - Efforts to manipulate behaviour are ancient, but they may gain a new quality when they use AI systems
  - Given users' intense interaction with data systems and the deep knowledge about individuals, they are vulnerable to manipulation, and deception.
  - Many advertisers, marketers, and online sellers will use any legal means at their disposal to maximise profit, including exploitation of behavioural biases, deception
  - Improved AI "faking" technologies make what once was reliable evidence into unreliable evidence. E.g Deep fake

"Privacy protection has diminished massively compared to the pre-digital age when communication was based on letters, analogue telephone communications, and personal conversation"

# Ethics of AI and Robotics

- Human-Robot Interaction
  - Human-robot interaction (HRI) now pays significant attention to ethical matters
  - Human-human interaction has aspects that appear specifically human in ways that can perhaps not be replaced by robots: care, love, and sex
  - Examples:
    - Care robots: the use of care robots ….. future of de-humanised care
    - Sex robots: are a continuation of slavery and prostitution