Robotics Topic 7 – Robots, AI and ethics

Prepared for the module team by Daniel Gooch

1 Robots, AI and ethics

Welcome to Topic 7.

In this topic you will explore how ethical choices within robotic and AI systems have a significant impact on whether they are socially beneficial or not. The key message from this topic is that just because we *can* build something that is technically amazing, it doesn't mean we *should*. We will start with a brief summary of different ethical frameworks before using real-world case studies to illustrate some of the limitations of machine learning. We will briefly explore the complex legal and ethical implications behind autonomous cars. The major issue that is considered throughout this block is the relationship between humans, society and robotic and AI systems, and how we can best use robots to the betterment of society.

1.1 Learning outcomes

By the end of Topic 7 you should be able to:

- understand three different perspectives on ethics
- appreciate the particular challenges associated with robots, AI algorithms and ethics
- understand the difficulties of developing robots that have an ethical approach to decision making
- argue whether you believe that autonomous cars can act ethically.

2 What are ethics?

As robotic and AI systems are increasingly integrated into our daily lives, they are increasingly discussed in terms of the ethics they hold. Is the use of robotics in warfare ethical? What ethical standards should drive the decision making of autonomous cars? What are the ethical shortcomings of many machine learning systems? Before delving into the complexities of these discussions, let us take a step back and consider this question: what is meant when we talk about ethics?

Ethics, or moral philosophy, is a branch of philosophy that considers the nature of right and wrong and can help support decision making. Ethics can operate at many different levels: an individual's moral code, the ethical judgement of an organisation, or the decision making of a national or international government.

Ethical standards should not be conflated with legal standards. The law specifies activities that societies have determined should be punished; for example, you should not steal a car. However, ethics is concerned with the correctness of an activity and might actually conflict with the law. As an example, if you stole a car to thwart a terrorist attack, then you will have broken the law (it is illegal to steal a car), but many people would believe you acted ethically (by stopping terrorism). On the other hand, tax avoidance is a legal strategy used by individuals and corporations to reduce the amount of tax paid in a particular country; however, it is increasingly seen as being ethically dubious since it deprives those countries of income needed to pay for basic services.

What makes this so complicated when it comes to defining how robotic and AI systems should behave is that not only do different people have different perspectives on what behaviours are ethical, they may believe in completely different ethical frameworks. A framework is a set of principles and values used by a person to guide their decision making.

In the remainder of Section 2 we will very briefly explore three different ethical frameworks. Don't be too concerned about the terminology, which is technical in nature and rarely used: your focus should be on understanding why embedding ethics into robotic and AI systems is so challenging.

2.1 Virtue ethics

Virtue ethics originate in ancient Greek philosophy, specifically Socrates (c. 469–399 BC). This tradition emphasises that an ethical individual will have developed positive characteristics (i.e. virtues) which will make that individual habitually act in an ethical manner.

Aristotle argued that we acquire these virtues by repeated actions and corrections. In his famous work, *Nicomachean Ethics*, Aristotle discusses 11 key moral *virtues*, including:

- **courage** in the face of fear
- **truthfulness** with self-expression
- **friendliness** in social conduct.

Each of these virtues sits between two corresponding *vices*: one of excess and one of deficiency. For example, too much courage is actually *rashness*, whereas too little is *cowardice*.

While virtue ethics remained popular in Europe throughout the Middle Ages, it declined during the seventeenth and eighteenth centuries with the development of alternative frameworks. However, virtue ethics has regained popularity with some modern philosophers who believe that in abandoning it we have lost an understanding of how to develop moral characters.

From the perspective of embedding ethics into robotic and AI systems, virtue ethics suffers from a challenge that you may have already noticed. By emphasising the value of the qualities of the individual making the decision, it doesn't focus on the sorts of actions which are morally permissible. In this context, it is therefore not necessarily a useful ethical framework.

2.2 Deontological ethics

The root of the word ‘deontological’ is *deon*, which is Greek for ‘duty’ or ‘obligation’. Deontological ethics is an approach that focuses on the correctness of the decision to undertake a particular action, regardless of the result of that action. Under this framing, the focus is on ensuring that the motivation, principles and ideals underpinning the decision to act are correct. Therefore, a person can perform an action that has terrible consequences so long as they had ethical motives.

Deontological ethics were first described in the eighteenth and early nineteenth centuries by philosophers including Immanuel Kant (1724–1804). A number of deontological philosophers have proposed rules of moral behaviour; an example is the list below from Bernard Gert (2004):

1. Don’t kill.
2. Don’t cause pain.
3. Don’t disable.
4. Don’t deprive of freedom.
5. Don’t deprive of pleasure.
6. Don’t deceive.
7. Keep your promise.
8. Don’t cheat.
9. Obey the law.
10. Do your duty.

Such an ethical framework is somewhat useful for embedding ethics into robotic and AI systems. In fact, Asimov’s Three Laws of Robotics are a set of deontological rules of behaviour. However, that doesn’t mean the challenges are over.

First of all, selecting a set of moral rules to embed in robotic and AI systems is not straightforward, as no one can conclusively state that a particular set of rules is (a) correct and (b) comprehensive enough to cover all scenarios. Secondly, many would argue that codifying the meaning of the concepts in the rules is not possible: while a person may have a clear understanding of ‘pleasure’, codifying that for a machine is likely to be impossible. Returning to Asimov’s rules, it should be remembered that Asimov was a fiction writer interested in telling stories. His Laws of Robotics were designed to help tell these stories: Asimov did not need to worry about how – or indeed if – they could ever be implemented.

Finally, there is always the problem of what to do when ethical rules conflict. Asimov deals with this by setting a priority level for each of the rules. Whether that would work in practice very much remains open for debate. For example, the cartoon in Figure 2.1 illustrates the deontological approach to answering ‘Do you not like the birthday present I got you?’, prioritising not deceiving over causing pain and depriving pleasure.

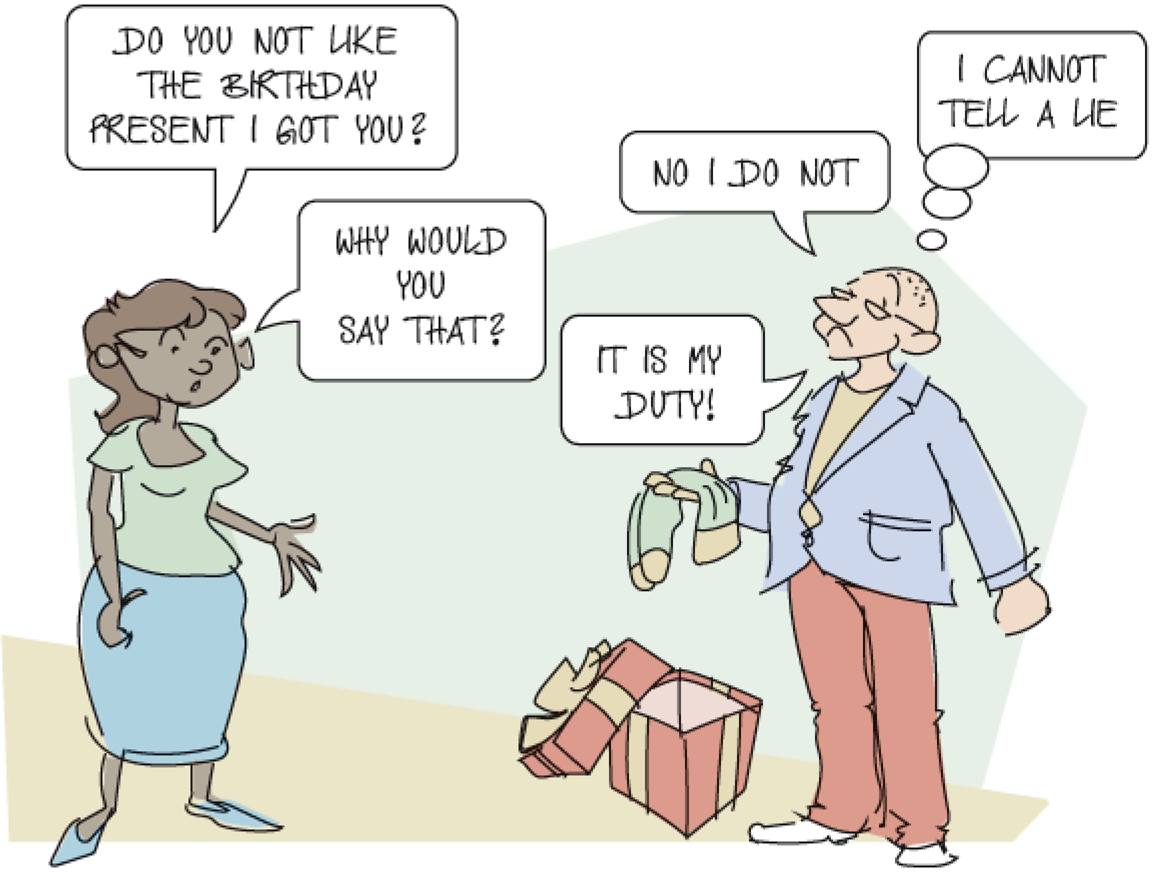

Figure 2.1 A cartoon illustrating the deontological response to whether you should answer a question truthfully

Hide description

A cartoon figure of a man and a woman. The lady is asking the man ‘Do you not like the birthday present I got you?’ who thinks ‘I cannot tell a lie’ and replies to the woman ‘No I do not’.

The woman is upset and asks ‘Why would you say that?!’ to which the man replies ‘It is my duty!’

While a trivial example, imagine a search and rescue robot who has found a victim of a building collapse. It has calculated that it can only save that person by driving over the legs of a different person who is trapped, breaking their legs. The robot is programmed with Gert’s list of rules. Does the rule of ‘don’t kill’ override that of ‘don’t cause pain’? Do you think there are circumstances where the ‘don’t cause pain’ rule overrides that of ‘don’t kill’?

2.3 Consequentialist ethics

The final framework we will consider is known as consequentialist ethics. This approach focuses on the result of the action: if an action has a ‘good’ outcome, then the action was the correct one to take. This is commonly expressed by the phrase ‘the end justifies the means’. Under this framing, a person can perform terrible actions as long as the overall consequences are positive.

Although this framework has its origins in a strand of Ancient Greek philosophy, it developed during the eighteenth and nineteenth centuries under people such as Jeremy Bentham (1748–1832) and continues to influence the modern world through people such as the hugely-influential economist Milton Friedman (1912–2006).

One of the more challenging aspects of this framework is that determining the intended and unintended consequences of an action may be extremely difficult. An additional challenge is assessing the respective 'goodness' of each consequence, especially when different consequences interact.

For example, the cartoon in Figure 2.2 illustrates the consequentialist approach to answering the question 'Do you not like the birthday present I got you?' and justifies not upsetting the lady by telling a small lie.

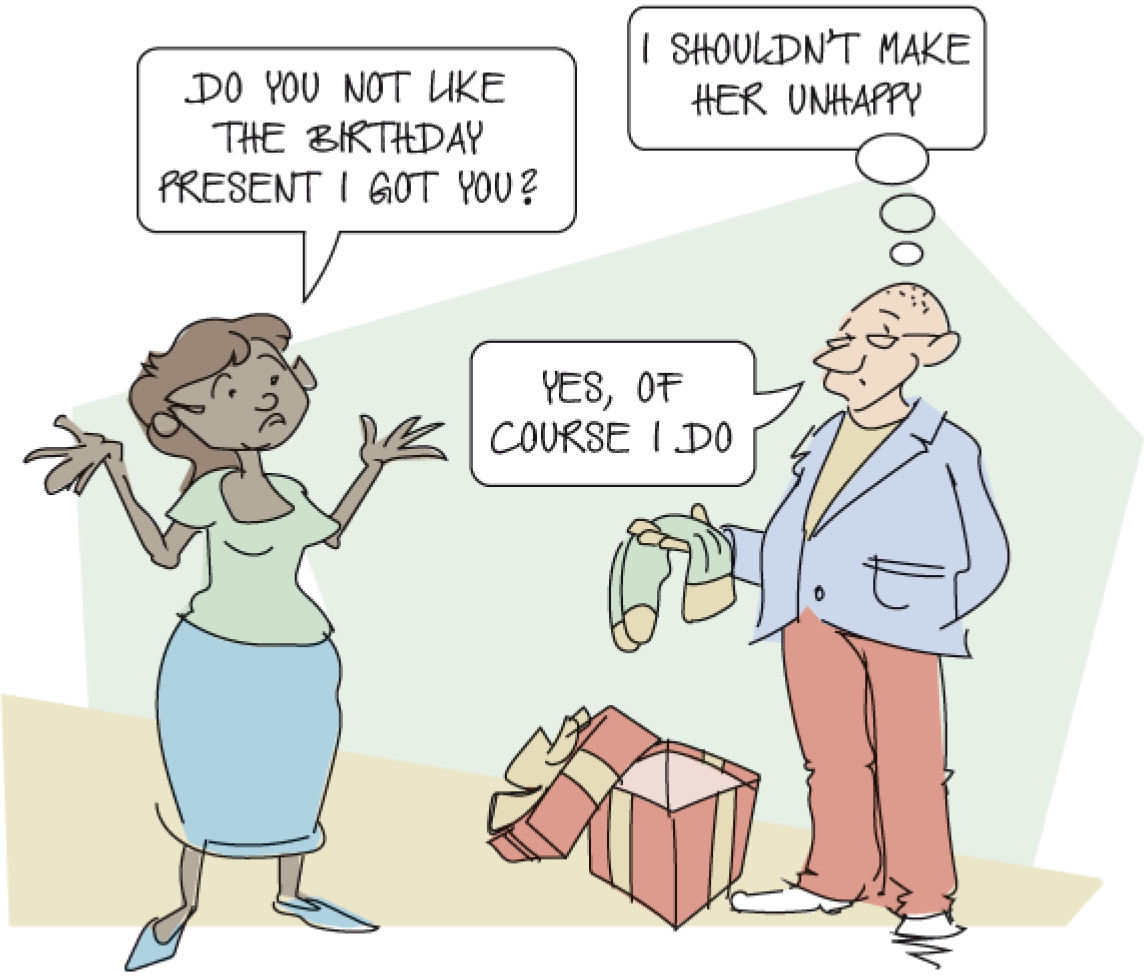

Figure 2.2 A cartoon illustrating the consequentialist response to whether you should answer a question truthfully

Hide description

A cartoon figure of a woman and a man. The woman is asking the man 'Do you not like the birthday present I got you?'. The man thinks 'I shouldn't make her unhappy' so replies 'Yes, of course I do'.

This again is a trivial example, where the consequences are small regardless of the decision of whether to lie or not. Now imagine a delivery robot that has been used to provide life-saving, but illegal, painkillers to a person who is extremely unwell. It has been caught by the police and is being questioned. What are the consequences of it telling the truth to the police? What is the correct course of action for the robot?

Activity 2.1 Discuss

Robots and AI systems are being increasingly used in the healthcare profession. It is not inconceivable that in the next 10 years, we will have robotic assistants monitoring a ward of patients.

Spend five to ten minutes thinking about which ethical framework you would choose to embed within such a robot. You might like to think about:

a. the outcomes of poor decision making
b. practicalities of encoding the ethics into the robot.

Discuss your ideas with others in your Cluster group forum.

3 The ethics of machine learning

In Topic 2 we spent some time exploring neural networks and how these underpin progress in the field of machine learning. Machine learning algorithms now underpin much of our daily lives, and can be found in systems including email spam detection, image recognition, automated language translation and systems determining which job applicants a company should hire.

Genuine progress has been made in machine learning processes over recent years; however, it has also become a convenient label for corporations and governments to sell products and policies without sufficient care being taken to understand the actual limits of machine learning.

Figure 3.1 A tweet from Professor Arvind Narayanan

Hide description

A tweet from Professor Arvind Narayanan

Key point 1. AI is an umbrella term for a set of loosely related technologies. Some of those technologies have made genuine, remarkable, and widely-publicised progress recently. But companies exploit public confusion by slapping the 'AI' label on whatever they're selling.

The computer scientist Professor Arvind Narayanan from Princeton University has developed a grouping of machine learning algorithms in terms of their success.

1. The first grouping is *perception tasks*, including activities such as image recognition, facial recognition, speech to text conversion and medical diagnoses from scans (see Figure 3.2). Arguably, machine learning algorithms are already at a point where they are as good as humans at these tasks, partially because the success criteria can be made clear. This makes training the systems much more straightforward.

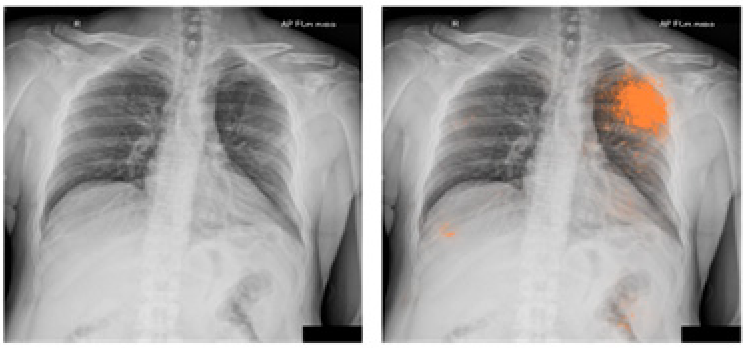

Figure 3.2 A chest X-ray depicting an issue of concern identified by machine learning but missed by an individual radiologist

Hide description

Two chest X-rays side by side. The X-ray on the left has not been adapted. The X-ray on the right has a cluster of orange marks in the upper right of the lung, highlighting the important regions identified by the machine learning algorithm.

2. The second grouping is *automating judgement* and involves systems such as spam detection, content recommendation and detection of copyright material. These systems are imperfect but are often 'fit for purpose'. Machine learning will never be perfect at such activities, because reasonable people can disagree about the correct decision. With sufficient systems for people to seek a change in the decision – for example, that an email shouldn't have been listed as spam – then in some circumstances, these algorithms work sufficiently well for us to have integrated them usefully

into our daily lives. There are limitations, where these algorithms are used in inappropriate contexts – for example, there is some evidence that automated sentencing decision making in the USA is racially biased (Thompson, 2019).

3. The third grouping is predicting social outcomes. This includes tasks such as predicting criminal reoffending, predicting which areas should be policed, and hiring workers. This area is fundamentally dubious, in part because we cannot predict the future – at least not well enough to provide the accurate feedback that machine learning algorithms need to become useful. This is the grouping where Professor Narayanan argues we are currently seeing a huge increase in AI ‘snake oil’, with organisations promising outcomes they cannot possibly deliver (see Figure 3.3).

Figure 3.3 Another tweet from Professor Arvind Narayanan

Hide description

A tweet from Professor Arvind Narayanan

Key point 2: Many dubious applications of AI involve predicting social outcomes: who will succeed at a job, which kids will drop out, etc. We can’t predict the future – that should be common sense. But we seem to have decided to suspend common sense when ‘AI’ is involved.

As the majority of autonomous robots use some form of AI system and machine learning as their sources of intelligence, we need to consider the ethical limitations of these systems.

3.1 Data biases

Machine learning algorithms are only ever as good as the data used to train them. This is sometimes reflected in the phrase ‘rubbish in, rubbish out’, a phrase from the field of computer science which highlights that flawed input data will produce nonsense outputs. This can have some interesting implications for machine learning algorithms. The following three sections consider three examples.

3.1.1 What is a fish?

Tench are fish popular with anglers. In refining an image recognition machine learning algorithm, a pair of researchers decided to incorporate anglers’ pictures of tench into their data set.

When they examined what features of tench their algorithm was focusing on, they found it was typically highlighting fingers with a green background, as shown in Figure 3.4.

Figure 3.4 Feature images from the tench recognition algorithm

Hide description

A figure made up of 14 small images in two rows. Each of the small images contains a small number of fingers and a green background.

Fish are not typically known for having fingers, so this was a little surprising! However, since tench are a trophy fish, when someone catches one, they tend to photograph it being held like a trophy (see Figure 3.5).

Figure 3.5 A tench being held like a trophy

Hide description

A fish being held like a trophy. The fingers of the person holding the fish are clearly visible.

There is no fundamental reason why this would be problematic – if you want to recognise trophy images of tenches. However, let us imagine that the researchers hadn't explored what features were being identified, and simply stated that their algorithm was excellent for identifying tench. Now imagine that their algorithm was added to an autonomous underwater vehicle tasked with monitoring tench in their native environment. It is unlikely that such a vehicle would be effective at its job.

3.1.2 Can a computer be sexist?

The problem with data biases goes beyond simple recognition problems: more worryingly, it can reflect the biases present in society. A widely reported case was that of Amazon's AI recruiting tool used to automatically process job applications. However, it was later discovered that this system had serious shortcomings.

Amazon's recruiting tool had been trained on a data set made up from the previous 10-years' worth of hiring decisions. Since the technology industry is highly male-dominated, the vast majority of people hired in that period had been men. This bias in the data was interpreted by the learning algorithm in a manner that preferred male candidates: the algorithm had inadvertently been programmed to be sexist.

According to the Reuters article which broke the story:

> It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges, according to people familiar with the matter.
>
> (Dastin, 2018)

A further problem was in the criteria it used to choose candidates. Most applicants listed typical skills you would expect from computing professionals, such as writing code. Since these were very common, the algorithms weighted those features as being unimportant. Instead, rare key words such as 'executed' or 'captured' were rated as being highly significant – despite having little to do with whether the candidate would be good at doing the job. This meant that not only were the algorithms hugely biased towards male candidates, they also ended up recommending unqualified candidates for all manner of jobs.

3.1.3 Can a computer be racist?

As a final example, we will consider Tay, a chatbot released by Microsoft in 2016 intended to interact with 18–24-year-olds on social media, particularly Twitter. As a chatbot, Tay was designed to learn how to converse naturally in a tone of voice similar to the people it was interacting with.

Within 24 hours, Tay had started posting sexist, anti-Semitic and racists tweets. Microsoft claim that it resulted from 'a coordinated attack by a subset of people [who] exploited a vulnerability in Tay' (Lee, 2016). Essentially, Twitter trolls bombarded Tay with so many sexist, racist tweets that Tay mistakenly learned that this was an acceptable and commonplace way of speaking.

> As many of you know by now, on Wednesday we launched a chatbot called Tay. We are deeply sorry for the unintended offensive and hurtful tweets from Tay, which do not represent who we are or what we stand for, nor how we designed Tay. Tay is now offline and we'll look to bring Tay back only when we are confident we can better anticipate malicious intent that conflicts with our principles and values.
>
> *Peter Lee, Corporate Vice President, Microsoft Healthcare, March 25, 2016*
>
> (Lee, 2016)

Given the prevalence and ease of finding unacceptable views on Twitter, it is perhaps of no surprise that, when let into the wild, machine learning systems also learn these unacceptable views.

Study note – More on Tay

Whilst Tay's tweets have been removed, it is still possible to find images on Twitter of some of these messages. If you want to see some examples, be warned that they are extremely offensive and may be distressing.

SAQ 3.1

Speech recognition systems are reliant on huge numbers of audio recordings for training purposes. What potential biases do you think this could have on speech recognition systems?

Hide answer

Answer

There are many different ways in which speech recognition system could be biased. Some voice-related issues include:

- regional accents
- national accents
- gendered voices
- the age of the speaker
- speech impediments.

A further concern could be that the behaviour of the speech recognition system changes our speech patterns as a society. How often do users say please or thank you to their Alexa or Google Home device?

There is also concern around the development of voice assistants and how their design can lead to poor 'treatment' by their users. One of the most common concerns is the predominant use of female voices acting in subservient ways, which could lead to the reinforcement of gender stereotyping. For a good summary article on this issue, read 'Artificial Intelligence has a gender bias problem – just ask Siri' on *The Conversation* (Adams, 2019).

3.2 Ethical concerns of AI algorithms predicting social outcomes

There are many concerns with the ethics of machine learning algorithms that go beyond the biases of input data. In this section we will draw on the example of the assessment of teachers to illustrate some of the difficulties of using machine learning algorithms in predicting social outcomes.

Study note – Further reading

If you are interested in the ethical dilemmas raised by machine learning algorithms, there are many different books, articles and programmes you can watch.

We recommend the book *Weapons of Math Destruction* by Cathy O'Neil (2016) as a good starting point. While somewhat USA-centric, the book provides an excellent step-by-step consideration of some of the problems that machine learning algorithms can cause when applied to unsuitable problems. The teaching example we are using to illustrate this section is included in this book, as well as in *an article in The Washington Post* (Turque, 2012).

Teaching is a complicated profession: while there are many aspects of education that educators understand and can control, results for students remain unpredictable and result from a huge range of factors.

Despite the complexity of education, recent years have seen a move in both the UK and USA education systems to try to evaluate teacher performance. A common evaluation is through so-called *value-added scores* where algorithms attempt to determine whether a teacher has improved the learning of their students by a desired amount. The effectiveness of this process has been, and continues to be, questioned (e.g. Morris *et al*., 2018).

We will focus on the story of Sarah Wysocki, a teacher in the District of Columbia (DC), USA. Whilst Sarah was well regarded by her colleagues and principal, she was fired because her students fell short of their predicted outcomes, meaning she had a low value-added score. DC had a long history of underachievement in its school system and the city government employed a contractor (Mathematica Policy Research) to develop an automated teacher evaluation system. The system would supposedly identify inadequate teachers in order that they could be fired and replaced with better teachers.

In theory the algorithm would use test scores to make an impartial, unemotional decision on whether particular teachers were effective at their job. Delegating the task to machines, in theory, removes potential biases caused by particular teachers being especially likeable (or indeed unlikeable), or those who have well-developed links with the community. However, the form the algorithm takes, and the response to Sarah's appeal, highlights a number of classic ethical concerns with machine learning systems.

3.2.1 Reducing complex scenarios to too few quantifiable parts

As we've already noted, teaching is a complicated process. There are a huge host of factors outside of a teacher's control that have an impact on students. Some long-term factors are poverty, learning disabilities and bullying; short-term factors include random incidents such as illness, crime or a family emergency. Any of these factors can affect how a student performs on a single test on a particular day. One-third of all adults in Washington DC, where Sarah taught, are considered 'functionally illiterate'; the city has the highest poverty rate and the highest rate of child hunger in the USA.

While a test score can easily be compared to expected performance, is that score an accurate reflection of the quality of the teacher?

3.2.2 Lack of feedback to refine the algorithm

Machine learning algorithms are trained using a feedback loop to determine whether or not the algorithm is making correct decisions. If the output matches what is expected from the outcome, then the algorithm is considered trained; if it is not, then adjustments are made and the training process is repeated until a satisfactory result is produced. In the image prediction scenario, this is straightforward: it is easy enough to tell the algorithm whether or not a picture it has identified as containing a fish actually contains a fish. In the context of teaching evaluations, this feedback loop is missing, as how can you determine whether the given score for a teacher is correct or not?

Furthermore, the local school authority considered the algorithm's decision to be final. When Sarah took her case to appeal, the panel stated that the accuracy or not of the algorithm's score lay outside of their remit and they could not consider it. This leads on to our next concern.

3.2.3 'Black box', inexplicable algorithms

The algorithm that determined Sarah's score was secret. Despite asking a large number of individuals in positions of power, she could not find out how her score had been calculated – the algorithm was an impenetrable black box, with data going in, and scores coming out. If it is impossible to determine how decisions are made, then it is much harder to argue against the decision.

A common explanation for why algorithms cannot be disclosed is that they are commercially valuable and thus must remain secret due to commercial sensitivity. You will probably have your own views on whether this leads to more accurate algorithms or not.

3.2.4 Behaviour change to suit the algorithm

After much investigation, Sarah determined that many of her students arrived from a different school with potentially inflated test scores from the previous year. These students' test scripts in the previous year had an abnormally large number of corrections on them. This could be interpreted as a sign of cheating. If her students' incoming test scores were artificially high, this could explain why the difference between their predicted scores and actual scores were so large.

Sarah's case shows that when confronted with the possibility of automating decision making, the school authorities delegated their authority to the machine and did not question its decision.

While we may never know the truth behind Sarah's case, it does help illustrate that when faced with a system of measurement, some people will change their behaviours, or even cheat, to maximise their score within that system.

3.3 Summary

You may have noted that most of these concerns are a mixture of technical concerns and their application within society. This is undoubtedly where the majority of ethical concerns arise, where good intentions can have negative consequences.

4 Ethics and autonomous cars

Given their potential to fundamentally change the way many of us live our lives, it is worth considering the current state of the art of autonomous cars, and the ethical questions arising from these AI-driven robots.

4.1 What is an autonomous car?

Defining an autonomous car is not straightforward, as the term not only means different things to different people, but car manufacturers have their own definitions of autonomy. Generally, an autonomous car is regarded as a vehicle which uses a variety of sensors – including cameras, radar, lidar and GPS – as well as route planning and intelligence systems to move the car safely with little or no human input.

A host of companies have developed a range of approaches to autonomy, some more ambitious than others. Some companies, including the majority of car manufacturers, have opted to develop 'driver-assistive technologies' which automate some aspects of driving, whilst others are aiming for full autonomy. As Table 4.1 shows, autonomy can be divided into different levels. The most common classification comes from the US Society of Automotive Engineers (SAE) and ranges from no autonomy (Level 0) through to fully autonomous vehicles (Level 5). At the time of writing, no commercially-available vehicle has demonstrated Level 5.

While there continues to be work focusing on autonomous trucks and other vehicles, we will focus on cars because they help illustrate many of the ethical issues with autonomous vehicles. We will also focus on data from the USA as that is where many of the companies researching and selling cars with autonomous features are based, and where most trials of the technology have been conducted.

Show description

Table 4.1 A table detailing the Society of Automotive Engineers levels of driving automation (SAE

International, 2018)

	Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
What does the human driver have to do?	The person in the driver's seat is considered to be driving even when the support features are engaged.			You are not driving when these automated features are engaged even if you are in the driver's seat.		
	The driver must supervise the support features and is responsible for steering, acceleration, braking, etc.			When the feature requests, you must drive.	These automated features will not require you to take over driving.	
	These are driver support features			*These are automated driving features*		
What do these features do?	Provide warnings and momentary assistance to the driver.	Provide steering or brake and acceleration support to the driver.	Provide steering and brake and acceleration support to the driver.	Can control the vehicle under limited circumstances. Will not operate unless *all* required conditions are met.		Can control the vehicle under all circumstances
Example features	• Automatic emergency braking; • Blind spot warning; • Lane departure warning.	• Lane centring; or • Adaptive cruise control.	• Lane centring; and • Adaptive cruise control at the same time.	• Traffic jam chauffeur.	• Local driverless taxi services; • Pedals or steering wheel may not be installed.	• As with Level 4, but the feature is able to drive everywhere in all conditions.

4.2 The supposed benefits of autonomy

Why are companies so focused on developing the robots we call autonomous cars? There are two obvious benefits: improved safety and reduction in time spent driving.

1. **Safety**

   According to the World Health Organization (WHO, 2020), road traffic injuries are the leading cause of death for children and young adults aged 5–29 years, with more than 1.35 million people being killed each year. The most recent statistics for the UK show there were 1770 deaths and 165,100 casualties in 2018 (Department for Transport, 2019). The contributory factors behind these accidents have also been recorded:

   a. ‘injudicious actions’ (such as disobeying signs or speed limits) account for 20% of all accidents
   b. ‘impairment or distraction’ (drink-driving, drug use, tiredness etc.) 15%
   c. ‘driver error’ for 67%.

   You may have noticed that these three factors total more than 100%; this is because some accidents have more than one contributing factor.

   Whilst the rate of deaths and injuries has been falling (at least in developed countries), they are still high and, as the statistics suggest, mostly caused by human failings. By giving the control of vehicles over to computers and algorithms, we could perhaps further reduce the number of road traffic accidents.

2. **Time**

   Many people spend a lot of time driving. Data from 2017 revealed that the typical American car user spends 52 minutes per day driving – an increase of 4 minutes since 2014. In theory, autonomous cars will help generate more free time for people.

4.3 The current state of autonomous cars

It is challenging to distinguish between what companies say and what the reality is on the ground. However, there are some things that we know for certain. California has a reporting system whereby all licensed operators of autonomous cars must report the number of miles driven, and the number of disengagements (defined as the number of times a human has to take over). California also requires operators to report all traffic collisions involving an autonomous vehicle.

Using the most recent figures from 2018, Waymo (see Figure 4.1) drove the most miles at 1.2 million, followed by GM Cruise logging nearly 448,000 miles while Apple came in third at 80,000 miles. It is worth considering that in terms of complexity, driving a mile on a quiet motorway is very different from driving a mile around a busy city centre.

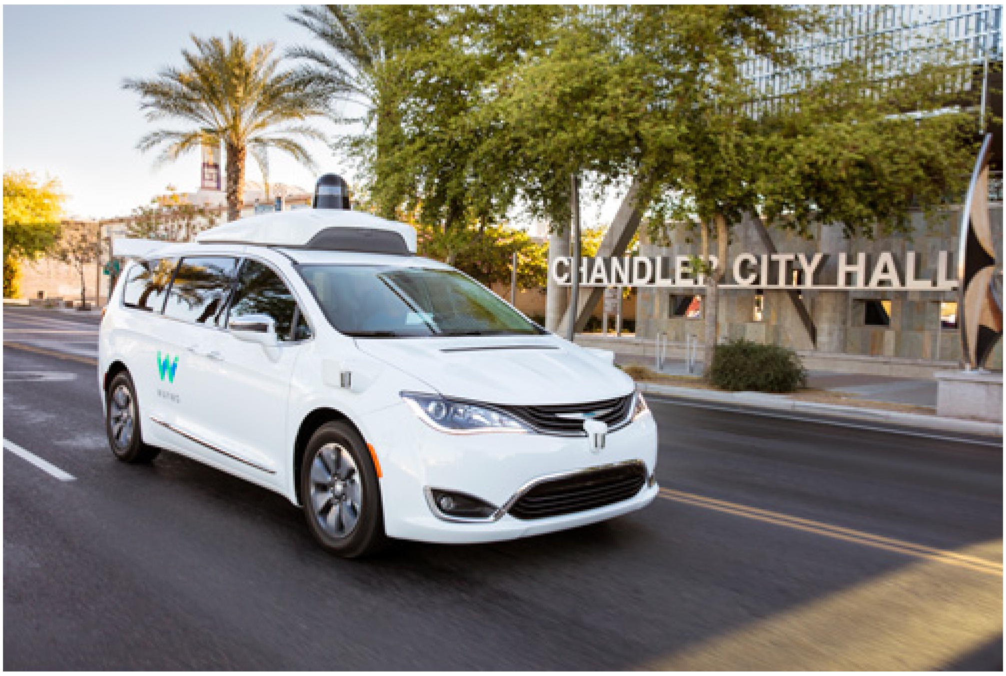

Figure 4.1 Waymo’s Chrysler Pacifica Hybrid minivan on public roads

Show description

Nuro, who reported 25,000 miles of driving, were upfront about some of the problems their vehicles experienced including:

- the cars can have difficulties in identifying objects
- the map information the cars rely on can be out of date and/or inaccurate
- the cars can make bad decisions (e.g. ‘planned trajectory resulted in erroneous sharp braking; recklessly tailgating motorist may have been unable to stop.’)

In addition to research vehicles, some manufacturers have released driver-assistive cars. Tesla is the most well-known of these, with the Autopilot mode providing some level 3 autonomy features.

However, these cars are not foolproof and, amongst other issues, can be susceptible to misuse. An up-to-date case study from the *MIT Technology Review* (O’Neill, 2020) has demonstrated how sticking a two-inch long piece of black tape to a road sign tricked multiple Teslas into accelerating to beyond the speed limit. The sign was 35 but was modified to read 85. The car could read the sign, but it could not infer the wider context of knowing whether such a speed was appropriate for the location.

Activity 4.1 Watch

Watch the following clip of the attack.

Recently some companies have been given licences to operate autonomous passenger services. For example, Waymo has a licence in constrained locations in both Arizona and California.

4.4 Ethical questions surrounding autonomous vehicles

Allowing autonomous passenger services raises our first ethical issue. Whilst public testing of these autonomous vehicles is legal, and passengers have consented to travel within that vehicle, everyone in the vicinity of that experimental car is an unwitting participant without consent. Is that ethically justifiable?

Why might this be problematic? Watch the video in Activity 4.2, which outlines a classical thought problem from the field of ethics – the trolley problem.

Activity 4.2 Watch

Watch the following explanation of the trolley problem.

The trolley problem is relevant to autonomous cars as it represents an ethical dilemma that might confront a vehicle. Road situations are complicated, particularly in busy cities with pedestrians, cyclists, and drivers of buses, cars and lorries all making split-second judgements, some of which will be incorrect. To operate in this world, sooner or later an automated car will have to make a decision akin to that outlined in the trolley problem – perhaps whether to crash itself and kill the rider, or to kill a group of pedestrians.

4.4.1 Autonomous car fatalities

The trolley problem is not just theoretical: the first road death from an autonomous vehicle has already occurred. Elaine Herzberg was killed on 18 March 2018 while crossing the road with her bicycle in Tempe, Arizona (Lee, 2019). She was hit by an Uber vehicle operating in self-drive mode with a human safety backup driver sitting in the driving seat. After the incident, Uber temporarily suspended all trials of its self-driving cars.

Activity 4.3 Listen

Listen to the following two clips from the BBC World Service programme *Robots on the road*.

Any road traffic death is a tragedy. This particular case is especially significant as it helps demonstrate many of the ethical shortcomings of autonomous vehicles.

In terms of safety, this incident demonstrates that autonomous cars are not a panacea: people will continue to be killed on the roads. It also puts into context the number of miles being driven by autonomous vehicles: the top three companies we reported from California total around 1.7 million miles. The national statistics in the USA show that the number of fatalities per million miles driven in 2017 was 1.16 (NHTSA, no date). Over 3212 *billion* miles of road use were recorded for that year. Autonomous cars are simply not travelling the volume or complexity of miles to establish their actual level of safety.

4.4.2 Autonomous cars and the law

One of the main ethical questions highlighted by the tragic case of Elaine Herzberg is that of liability. At the moment, the law is lagging behind technical innovation. Focusing on the UK, there are two main questions that apply:

1. What level of autonomy are we discussing?
2. Who is held responsible for any accident?

Should a fatal accident be treated as an insurance issue, a commercial accident, or manslaughter?

In a situation where a human is driving a non-autonomous car, the situation is clear. By driving, we accept the risk of being injured or killed. However, we expect that if an individual does cause harm, then they will be punished through the justice system.

The legal system for autonomous vehicles is still being developed. The UK Parliament has passed the Automated and Electric Vehicles Act 2018, although, at the time of writing, it has not yet been tested in court. Under the Act, an accident caused by an autonomous vehicle is the liability of the company who has insured that car – not the passengers, nor the developers of the car themselves. Under the Act, the insurer has the right to claim their costs from the ‘person responsible for [the] accident’, such as the manufacturer or software company who produced the code. Since this has yet to be tested in court, it is still unclear how effective this Act will be in regulating the industry, especially when autonomous vehicle developers may be based outside of the UK.

The Law Commission (2020) is continuing to review legislation around autonomous vehicles. This consultation is more concerned about the use of autonomous cars as a passenger-carrying service, rather than being used by the owner of the vehicle. The flow chart in Figure 4.2 shows the Law Commission’s 2019 proposal, where unlike the Automated and Electric Vehicles Act 2018, it would be the manufacturers who would generally be held responsible for accidents if it could be demonstrated they were at fault. However, how easy that will be to prove remains to be seen.

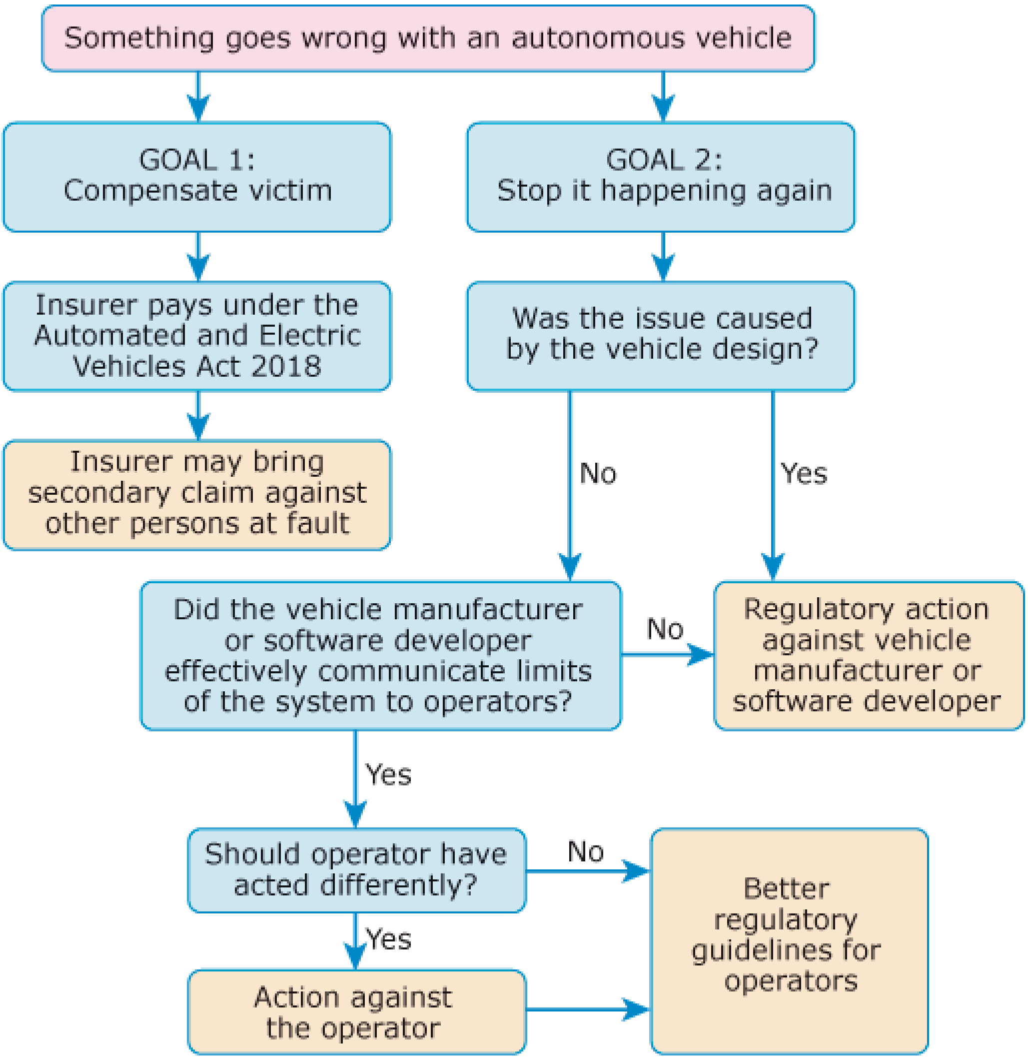

Figure 4.2 The Law Commission's proposal for legislation for passenger services using autonomous cars

Show description

ePortfolio

You can complete and submit the following activity as part of your ePortfolio.

Activity 4.4 Discuss

Consider your personal view of the ethics of autonomous vehicles.

a. Who do you think should be responsible for accidents caused by autonomous vehicles? Why?
b. How do you think autonomous vehicles could be made to behave in an ethical manner?

Share you views in your Cluster group forum.

4.5 The impact on society

According to statistics from the UK Department for Transport (2020), there are around 322,000 heavy goods vehicle drivers, 363,000 licensed taxi drivers and 114,000 bus and coach drivers in the UK. If we include delivery drivers, then there are likely to be more than a million jobs that are based around driving. Autonomous vehicles will potentially displace many – if not all – of these roles, which could result in mass unemployment. If you then add in all of the dependent industries – such as the vehicle insurance industry – you should begin to see how transformative the mass rollout of autonomous vehicles would be to the economy. The threat from automation on the labour market is beginning to attract political attention. The risk is not just to drivers: tasks involving supervising or operating machinery, routine sales jobs and those involving decision-making are most at risk. And not all threatened roles are low paid: financial advisors, insurance actuaries and mortgage brokers are all tasks ideally suited to computerisation.

As a final point, there are also concerns over the privacy risk from the mass use of autonomous vehicles. These vehicles may collect large amounts of personal information about users alongside detailed location tracking. In some ways this is not new: as long ago as 2011, General Motors attempted to change the user agreement in its OnStar in-car safety and connectivity system so that GM could share 'anonymised' driver information with third-parties. Public reaction was so strong that GM was forced to backtrack just one week later (General Motors, 2019).

Study note – Anonymisation

Anonymisation of data is any process that supposedly removes or obscures information in the data that could be used to identify an individual or group of people. It is extremely difficult to perform true anonymisation in a manner that data cannot be combined with other anonymised data to 'deanonymise' people.

One way of approaching the privacy implications of such technologies is by attempting to answer five data-related questions posed by Fagnant and Kockelman (2015):

1. Who should own or control the vehicle's data?
2. What types of data will be stored?
3. With whom will these data sets be shared?
4. In what ways will such data be made available?
5. And, for what ends will they be used?

Depending on the answers to these questions, autonomous vehicles could become the greatest surveillance tools in history, or they may become tools of freedom, changing our society for the better.

5 The ethics of military robots

No discussion on the ethics of robotics would be complete without a brief exploration of military robots, some of which have been designed to kill.

Some military robots have been designed to support activities that most people would agree are ethically sound. Obvious examples are mine-clearing and bomb disposal robots, such as TALON and PackBot. Such robots are not lethal: they are not designed to kill or injure people, but to save lives. However, if you mount a gun onto such platforms, as was proposed with TALON (with the armed version named SWORDS), then it becomes a lethal autonomous weapon – something many people have ethical concerns about.

5.1 Lethal autonomous weapons

Lethal autonomous weapons (LAWs) can be defined in many ways; we are going to focus on systems which can autonomously search for and identify targets, decide to engage an identified target and subsequently engage that target without the intervention of a human. Such robots are shrouded in secrecy, although it is commonly held that all major arms-exporting countries – including the USA, UK, Israel, Russia and China – are developing LAWs.

Why are militaries around the world interesting in LAWs? In part it is because robots have capabilities that human-operated systems do not: they can be engineered to be smaller, faster and lighter. AI systems can respond faster than human operators can, while robots do not get tired, feel anger or seek revenge. Furthermore, the deployment of robots – as we have seen for decades with the rapid rise in the use of remote-piloted drones – can preserve the lives of the country that is deploying the robot.

The falling cost and increasing commercial availability of robotic and drone technology, as well as off-the-shelf AI systems, mean that LAW technologies will soon no longer be the preserve of nation states. You don’t have to be a science fiction author or Hollywood script writer to imagine a future of drones being used for terrorist attacks.

5.1.1 A brief history of LAWs

Depending on how you interpret the definition, some of the first LAWs were *defensive* systems, intended to protect naval ships from inbound aircraft and missiles. For example, the US Phalanx close-in weapon system (CIWS) made by General Dynamics (see Figure 5.1) integrates a computer-controlled radar and a multi-barrelled gun to engage any inbound anti-ship missiles. CIWS is capable of searching for, detecting, tracking and engaging targets without human intervention.

Figure 5.1 A US Phalanx close-in weapon system being test fired

Show description

More recently, countries have been interested in developing and using *offensive* LAWs. One of the most discussed is the Super aEgis II, a sentry gun developed for deployment in the Demilitarised Zone between South and North Korea. Exact details are unknown, at least in part due to the sensitive nature of the project, but some reports suggest that the sentry gun can operate without any human instruction.

Another focus for ground-based forces has been the development of tele-operated and autonomous tanks. Russia and China are both suspected to be working on such systems. The Israeli-developed Guardium vehicle is already in use, although exact details of its capabilities are unknown.

Naval vehicles are also being revolutionised by AI systems and autonomy. Sea Hunter is an autonomous ship developed by the US Government research division DARPA to demonstrate the potential of unmanned vessels for submarine hunting and mine clearance. Its capabilities in terms of autonomy are obscured by national security concerns, but it is reported to be able to navigate autonomously for thousands of miles, albeit with human oversight.

Progress is also being made with airborne military drones. Northrop Grumman's X-47B drone made headlines as the first drone to autonomously take off from, and land on, an American aircraft carrier. It was also the first unmanned drone to autonomously refuel while in flight. Military researchers in the USA have also been exploring the tactics of how to use large numbers of small, low-cost drones to engage targets, as can be seen in the video in Activity 5.1.

Activity 5.1 Watch

Watch the following clip from the US Navy demonstrating the swarming behaviour of Perdix drones.

Israel Aerospace Industries was the first company in the world to design and deploy a so-called ‘loitering’ autonomous air-to-ground missile. When fired, the Harpy missile will autonomously fly to a defined area where it scans for targets such as anti-aircraft missile emplacements and radar stations. It is reportedly capable of remaining within the defined area for 9 hours, searching for targets to engage.

There is now an ever-expanding range of LAWs for both defensive and offensive use. The reasons behind their deployment are clear; but as you will see in the next section, the ethical concerns about their development and deployment are less clear.

5.2 Ethical concerns of LAWs

For the purposes of this discussion, we are setting aside the issue of whether war can ever be ethical to focus specifically on the ethical concerns of LAWs.

International law dictates that certain conditions must apply in order for a military response to be considered legal. These include: distinguishing between civilian and military targets; using proportionate force; a clear judgement on the military necessity of a particular engagement; and the need to accept the surrender of enemy combatants.

However, objections to LAWs go beyond what might be legal. Some opponents to LAWs argue that AI systems will never be able to make these judgements, and therefore should only be developed where a human makes the ethical judgements. While AI systems can be excellent at recognising targets, in certain conditions, they are unable to understand the context and ethical framing of the decision-making process. For example, the Catholic Church has made their concern about LAWs clear:

The Holy See recognizes the difficulties and resistance in foreseeing and fully understanding the implications of autonomous weapon systems. One thing is certainly clear: the development of LAWS will provide the capacity of altering irreversibly the nature of warfare, becoming even more inhumane, putting in question the humanity of our societies and, in any case, compelling all States to reassess their military capabilities.

…

A world in which autonomous systems are left to manage, rigidly or randomly, fundamental questions related to the lives of human beings and nations, would lead us imperceptibly to dehumanization and to a weakening of the bonds of a true and lasting fraternity of the human family.

Archbishop Ivan Jurkovič, Permanent Observer of the Holy See to the United Nations, speaking at the 2018 Group of Governmental Experts on Lethal Autonomous Weapons Systems (LAWS), Geneva, Switzerland, 9 April 2018.

(Fair, 2018)

The Campaign to Stop Killer Robots is a coalition of non-governmental organisations (NGOs), and is led by Mary Wareham and Isabelle Jones. They identify six clear problems with LAWs, including:

1. Fully autonomous weapons would decide who lives and dies, without further human intervention, which crosses a moral threshold. As machines, they would lack the inherently human characteristics such as compassion that are necessary to make complex ethical choices.
2. Replacing troops with machines could make the decision to go to war easier and shift the burden of conflict even further on to civilians. Fully autonomous weapons would make tragic mistakes with unanticipated consequences that could inflame tensions.
3. Fully autonomous weapons would lack the human judgment necessary to evaluate the proportionality of an attack, distinguish civilian from combatant, and abide by other core principles of the laws of war. History shows their use would not be limited to certain circumstances.
4. Fully autonomous weapons could be used in other circumstances outside of armed conflict, such as in border control and policing. They could be used to suppress protest and prop-up regimes. Force intended as non-lethal could still cause many deaths.
5. It’s unclear who, if anyone, could be held responsible for unlawful acts caused by a fully autonomous weapon: the programmer, manufacturer, commander, and machine itself. This accountability gap would make it is difficult to ensure justice, especially for victims.
6. The US, China, Israel, South Korea, Russia, and the UK are developing weapons systems with significant autonomy in the critical functions of selecting and attacking targets. If left unchecked the world could enter a destabilizing robotic arms race.

(Campaign to Stop Killer Robots, no date)

This summary, taken from the campaign's website, highlights some of the ethical concerns with LAWs. One issue that is not listed by the campaign is the potential damage that could be caused by people hacking into an autonomous weapon system.

The campaign is seeking to prohibit the development, production, and use of fully autonomous weapons through legislating a ban using national laws and international treaties. In 2015, more than 1000 scientists and engineers, including experts in robotics and artificial intelligence, signed a joint letter expressing their concerns:

> Just as most chemists and biologists have no interest in building chemical or biological weapons, most AI researchers have no interest in building AI weapons – and do not want others to tarnish their field by doing so, potentially creating a major public backlash against AI that curtails its future societal benefits. Indeed, chemists and biologists have broadly supported international agreements that have successfully prohibited chemical and biological weapons, just as most physicists supported the treaties banning space-based nuclear weapons and blinding laser weapons.
>
> In summary, we believe that AI has great potential to benefit humanity in many ways, and that the goal of the field should be to do so. Starting a military AI arms race is a bad idea, and should be prevented by a ban on offensive autonomous weapons beyond meaningful human control.
>
> *AUTONOMOUS WEAPONS: AN OPEN LETTER FROM AI & ROBOTICS RESEARCHERS*
>
> *24th International Joint Conference on Artificial Intelligence (IJCAI-15), July 2015.*
>
> (Future of Life Institute, 2015)

Amongst the signatories were the late Stephen Hawking, Elon Musk (founder of Tesla and SpaceX), Steve Wozniak (co-founder of Apple) and Demis Hassabis (CEO of Google DeepMind). The letter attracted a great deal of attention since it showed many leading technologists were concerned about LAWs.

In 2019, UN delegates met in Geneva to discuss potential restrictions on LAWs. The majority of member states, along with the UN Secretary General, called for either a total ban or strict regulation on the development and deployment of LAWs. The policy was blocked by a group of countries including the USA, UK, Russia, Israel and Australia. At the time of writing, there are no international conventions on the development and use of LAWs.

Activity 5.2 Discuss

Do you think there are any advantages to using autonomous weapons? What potential problems do you think autonomous weapons could cause? Would you be in favour of a ban?

> Discuss your ideas with others in your Cluster group forum.

6 Professional responsibilities

By now, you should recognise that the role of a computer professional – whether specialising in robotics and AI systems or some other domain – does not just mean that you should build whatever computing system you want. Computing takes place in the context of society as a whole. The many examples in this topic show that whilst many designers have the best of intentions, their actions may have unintended negative consequences.

Naturally, computing professionals should act within the law; but they should also operate within social frameworks set by society and respect certain ethical boundaries.

As a relatively new profession, computing does not have a single professional body within the UK to develop and enforce ethical standards. The British Computing Society, the Association for Computing Machinery and the Institute of Electrical and Electronics Engineers each have their own codes of conduct incorporating ethical values. Members of these organisations pledge to adhere to these codes in their professional activities.

Whilst these codes are each slightly different from one another, there are large areas of overlap.

Each code of conduct is a set of principles intended to guide rather than prescribe. The real world is messy and the application of ethics to specific professional settings can often be ambiguous. The codes are intended to aid professional judgement rather than be a comprehensive set of rules and regulations.

Each of the codes is concerned with the primacy of the good of society. For example, the IEEE Code of Ethics has as its first principle:

> to hold paramount the safety, health, and welfare of the public, to strive to comply with ethical design and sustainable development practices, and to disclose promptly factors that might endanger the public or the environment
>
> (IEEE, no date)

The codes of conduct have similar concerns to the issues of deontological rules of ethics discussed in Section 2.2. Parts of the codes may be ambiguous and different parts of a code can conflict with one another. Significantly, the codes do not aim to foresee every possible situation. In such cases, professionals are expected to behave professionally and use their best judgement.

SAQ 6.1

What potential problems do you think you might face in following a code of conduct?

Hide answer

Answer

There are many potential problems you might face in following a code of conduct, depending on the area you are working in. Issues could include the following.

- You might put your job at risk.
- You might come under pressure from clients to act in a way you deem to be unethical.
- Ethical codes need strong institutional backing to function effectively. You may come under pressure from management to cut corners when projects are running late, for example.
- Societal expectations can be different in different countries and clash with your ethical expectations.

What other problems did you identify?

7 Practical activities

Activity 7.1 Robot practical activities

Follow the instructions on this week of the study planner.

8 Summary

In this topic, we have explored the ethics of how robotic and AI systems have an impact on society.

Robots and AI systems pervade many elements of daily life. While great strides have been made in the technical development of robotics and AI systems, you should not be swayed by the novelty and excitement of them.

Deep and systemic questions remain as to how we can develop ethical systems that robots understand, and how we put robots to work for the betterment of society.

Through the example of autonomous cars, we have seen how ethical concerns can be multifaceted, with the right path ahead unknown. As is common with many areas of technology, our technical and engineering skills move faster than our legal and ethical understandings.

Regardless of the context, being a computer professional means that you should act in accordance with your best professional judgement. Others may disagree with your decisions, but you should try to act in an ethical manner at all times.

Where next

This is the end of Topic 7.

Topic 8 considers the possible future of robotics and of societies which use robots and other intelligent systems.

References

Adams, R. (2019) ‘Artificial Intelligence has a gender bias problem – just ask Siri’, *The Conversation*, 22 September. Available at: https://theconversation.com/artificial-intelligence-has-a-gender-bias-problem-just-ask-siri-123937 (Accessed: 30 June 2020).

Campaign to Stop Killer Robots (no date) ‘The problem’. Available at: https://www.stopkillerrobots.org/learn/#problem (Accessed: 16 May 2020).

Dastin, J. (2018) ‘Amazon scraps secret AI recruiting tool that showed bias against women’, *Reuters*, 10 October. Available at: https://uk.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUKKCN1MK08G (Accessed: 15 May 2020).

Department for Transport (2019) ‘Contributory factors for reported road accidents (RAS50)’, *Department of Transport*, 26 September. Available at: https://www.gov.uk/government/statistical-data-sets/ras50-contributory-factors (Accessed: 27 May 2020).

Department for Transport (2020) ‘Latest transport statistics’. Available at: https://www.gov.uk/government/organisations/department-for-transport/about/statistics#latest-transport-statistics (Accessed: 16 May 2020).

Fagnant, D.J. and Kockelman, K. (2015) ‘Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations’, *Transportation Research Part A: Policy and Practice*, 77, pp. 167–181. doi:10.1016/j.tra.2015.04.003.

Fair, J. (2018) ‘Archbishop Jurkovic Warns of Autonomous Weapons’, *Zenit*, 11 April. Available at: https://zenit.org/articles/archbishop-jurkovic-warns-of-autonomous-weapons/ (Accessed: 16 May 2020).

Future of Life Institute (2015) ‘Autonomous Weapons: an Open Letter from AI & Robotics Researchers’, 28 July. Available at: https://futureoflife.org/open-letter-autonomous-weapons#signatories (Accessed: 16 May 2020).

General Motors (2019) ‘OnStar Reverses Decision to Change Terms and Conditions’, 27 September. Available at: https://media.gm.com/media/us/en/gm/news.detail.html/content/Pages/news/us/en/2011/Sep/0927_onstar.html (Accessed: 27 May 2020).

Gert, B. (2004) *Common morality: Deciding what to do*. Oxford University Press.

IEEE (no date) ‘IEEE Code of Ethics’. Available at: https://www.ieee.org/about/corporate/governance/p7-8.html (Accessed: 16 May 2020).

Law Commission (2020) ‘Automated Vehicles’. Available at: https://www.lawcom.gov.uk/project/automated-vehicles/ (Accessed: 16 May 2020).

Lee, D. (2019) ‘Uber self-driving crash 'mostly caused by human error'’, *BBC News*, 20 November. Available at: https://www.bbc.co.uk/news/technology-50484172 (Accessed: 16 May 2020).

Lee, P. (2016) ‘Learning from Tay’s introduction’, *Official Microsoft Blog*, 25 March. Available at: https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/ (Accessed: 15 May 2020).

Morris, T.T., Davies, N.M., Dorling, D., Richmond, R.C. and Smith, G.D. (2018) ‘Testing the validity of value-added measures of educational progress with genetic data’, *British Educational Research Journal*, 44(5), pp. 725–747. doi:10.1002/berj.3466.

NHTSA (no date) ‘FARS Data Tables’, *United States Department of Transportation*. Available at: https://www-fars.nhtsa.dot.gov/Main/index.aspx (Accessed: 16 May 2020).

O’Neil, C. (2016) *Weapons of Math Destruction*. Penguin Random House.

O’Neill, P.H. (2020) ‘Hackers can trick a Tesla into accelerating by 50 miles per hour’, *MIT Technology Review*, 19 February. Available at: https://www.technologyreview.com/2020/02/19/868188/hackers-can-trick-a-tesla-into-accelerating-by-50-miles-per-hour/ (Accessed: 16 May 2020).

SAE International (2018) ‘Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles’, J3016_201806, 15 June. Available at: https://www.sae.org/standards/content/j3016_201806/ (Accessed: 16 May 2020).

Thompson, D. (2019) ‘Should We Be Afraid of AI in the Criminal-Justice System?’, *The Atlantic*, 20 June. Available at: https://www.theatlantic.com/ideas/archive/2019/06/should-we-be-afraid-of-ai-in-the-criminal-justice-system/592084/ (Accessed: 15 May 2020).

Turque, B. (2012) ‘“Creative … motivating” and fired’, *The Washington Post*, 6 March. Available at: https://www.washingtonpost.com/local/education/creative--motivating-and-fired/2012/02/04/gIQAwzZpvR_story.html (Accessed: 15 May 2020).

WHO (2020) ‘Road traffic injuries’, 7 February. Available at: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries (Accessed: 16 May 2020).

Further reading

Hevelke, A. and Nida-Rümelin, J. (2015) ‘Responsibility for crashes of autonomous vehicles: an ethical analysis’. *Science and engineering ethics*, 21(3), pp. 619–630. doi:10.1007/s11948-014-9565-5.

Scharre, P. (2018) *Army of None: Autonomous Weapons and the Future of War*. New York: W.W. Norton & Company.

Acknowledgements

Grateful acknowledgement is made to the following sources:

Figure 3.1: Taken from: https://twitter.com/random_walker/status/1196870650322919428?s=20

Figure 3.2: Taken from: https://pubs.rsna.org/doi/pdf/10.1148/radiol.2019191293

Figure 3.3: Taken from: https://twitter.com/random_walker/status/1196870836092882949?s=20

Figure 3.4: Brendel, W. and Bethge, M. (2019) Eberhard Karls University of Tübingen, Germany, Werner Reichardt Centre for Integrative Neuroscience, Tübingen, Germany Bernstein Center for Computational Neuroscience, Tübingen, Germany

Figure 3.5: © Peter Faretra / https://www.flickr.com/photos/findmelost/4988103538 This file is licensed under the Creative Commons Attribution-Noncommercial-ShareAlike Licence http://creativecommons.org/licenses/by-sa/2.0/

Figure 4.1: © 2019–2020 Waymo LLC

Figure 4.2: https://www.lawcom.gov.uk/project/automated-vehicles/ Reproduced under the terms of the OGL, www.nationalarchives.gov.uk/doc/open-government-licence

Figure 5.1: © US Federal Government