The Good Future: A Brief Introduction to Ethics

Ben Mitchell

Introduction

As you approach the field of ethics, it is important to begin by understanding what 'ethics' is, and what it is not. In particular, it is important to distinguish between *ethics* and *morality*.

- Morality is about making value judgements, delineating right from wrong and good from evil.
- Ethics is the study of systems of morality.

The distinction here may seem subtle, and many people may use these terms interchangeably in everyday life. However, the difference is an important one: teaching you about *morality* requires you to ascribe to the teacher's values. Teaching you about *ethics* only requires you to examine your own values, and offers you tools to help refine them.

The primary reason for studying 'applied' ethics is to learn how to make decisions that will lead to a future you are happy with. If we accept the premise that the future is not pre-determined and set in stone, then we are left with the question of how to get a "good future," and ensure we don't end up in a "bad future." Ethics gives us mental tools we can apply to help us answer this question.

There are many different systems of morality we can examine as ethicists without committing to one of them as being 'right' and the others as being 'wrong.' That said, we need not treat all systems of morality as being equally useful. In particular, philosophers prefer systems which are 'consistent,' meaning both that there are not internal contradictions and that moral judgements can be made by application of logical reasoning from a set of base assumptions, such that anyone who agrees with the base assumptions should be able to reach the same conclusion about the morality of a particular action.

In other words, we want to be able to explain *why* a certain act is good or bad, not merely declare that it is. It is entirely possible to have a system of morality in which the answer to the question, "why is this act considered bad," is in essence "because I said so" (or "because my father said so," or "my holy book said so," or any other variation). However, it is exceedingly difficult to convince anyone else of the correctness of our judgements in this case, since unless they agree with the legitimacy of the authority we received our morality from, they have no reason to take our word for it.

Thus, ethicists seek to find sets of underlying principles that we hope many people can agree with, and from which we can hopefully derive most of the judgments we feel attached to. It may be that you and I can not agree about which holy book is the correct one, but we may still be able to agree that human life holds intrinsic value, and from there we can derive moral judgements such as, "murder is bad," in a way we can both agree is correct.

Several Ethical Frameworks

There are many introductions to ethics that are aimed at philosophy students; that is not the primary audience this text is written for (though of course, philosophers are welcome!). Rather, my aim is to provide an introduction to the subject for those who are mostly interested in "applied ethics." For this reason, I will treat ethical and moral frameworks as tools to be applied to decision making, rather than purely as a subject of philosophical study. Thus, where a philosopher might attempt to persuade you that their own preferred ethical framework is the best one, I will rather present a variety of frameworks and suggest when each might be of use, and when it might not. I encourage the reader to view these frameworks as tools in a toolkit, and try to match the most appropriate tool to whatever problem is in front of them (as opposed to picking the one you like best and then using it on every problem like the proverbial carpenter whose only tool is a hammer).

Note also that I am oversimplifying things here; this is intended as an overview and a starting point, not a comprehensive explanation of every variant on a given theme. Each of these frameworks has been studied and written about by many philosophers over many years, and there has been a great deal of careful thought about them that I am not describing here. This is also by no means an exhaustive list of available ethical frameworks; I present these in large part because they represent viewpoints with which the reader likely already has some familiarity. There are many other ethical frameworks out there, and the reader is encouraged to learn more about them.

Virtue Ethics

The framework known as 'Virtue Ethics' goes all the way back to Aristotle, and its core has remained largely similar for the last several thousand years. The basic idea is that rather than viewing 'virtues' and 'sins' as being in direct opposition, or even independent of each other, 'virtues' are points of balance on a spectrum where 'sins' are at the extremes.

For example, if we consider 'anger,' this framework would suggest that having too much anger is bad, because it would mean getting enraged over things that aren't really important, and lashing out at people who don't deserve it (e.g. beating up your roommate for leaving a dirty sock on the floor). On the other hand, too *little* anger is *also* bad, because it would mean we tolerate or ignore things that are really bad, and we *ought* to get upset about (e.g. watching someone get attacked, and just walking past with a shrug).

According to this framework, a virtuous person is one who seeks a balance between these

extremes (sometimes called the 'golden mean,' though this phrase was not used by Aristotle). The framework also suggests that we must *work* to become virtuous through practice; we should always be examining our own actions to make sure they are not becoming too extreme in any direction.

Many of us find this framework to be appealing, since it corresponds well with our intuitions that many forms of excess are likely to be harmful. It tends to be useful in making small, day-to-day decisions, and helps encourage us to consider and refine our behavior over time.

Virtue ethics is explicit that the decisions must be to some degree individual; what is the right amount of something for me may not be the right amount for you (e.g. 'how much should I eat' can have different answers depending on size, athleticism, etc.). This can be reassuring, but it also means that disagreements are difficult to resolve; there is no truly objective measure of morality. Additionally, because it is focused on the individual, it is often difficult to apply to policies or decisions that are likely to have large-scale consequences.

Deontology

This ethical framework is most commonly associated with Immanuel Kant, and is sometimes referred to as Kantianism; it is also described as the ethics of Duty. It postulates that the moral value (i.e. goodness or badness) of an act is inherent in the act itself. In other words, an action is good or bad in and of itself, and we should be able to determine its goodness or badness by logical study of the action.

Note that the root of the term *deontology* is the ancient Greek "ontology", which is the field of philosophy concerned with the nature of being and that which is inherent and essential to a thing ("de" in ancient Greek is a particle). Importantly, it is *not* related to the Latin "deo" (the ablative of "deus"), which translates as "god." Thus, deontological ethics has nothing to do with deity or religion, and everything to do with the *essential nature* of actions.

The fundamental principle of deontology is that we have an innate duty to act in a moral fashion, and that the morality of an action is innate to the action. The logical analysis states that anything we wish to be a moral rule must be *universalizable*; that is, it must be applicable universally without reaching a logical contradiction. In other words, we must be able to imagine *everyone* performing such an act without resulting in nonsense.

A common example is the question of theft; for example, let us say you have a very nice pen that I am covetous of, and I want to steal it. In order for it to be permissible for me to take your property and declare it mine, theft must universalize, so it must also be permissible for *anyone* to take the property of anyone else. However, if this were actually the case it would effectively mean that the basic concept of "private property" was no longer valid. But, if there is no such thing as private property, I cannot take something that belongs to someone else, because nothing "belongs" to anyone! This is a logical contradiction; thus, it must *not* be morally permissible for me to steal your pen, no matter how much I might

covet it.

The other way that deontology is frequently characterized is in the use of the *means-ends* test, which goes something like this:

> *It is never permissible to treat humans as merely means to an end, it is only permissible to treat them as ends in-and-of themselves.*

While the two forms are intended to get at the same fundamental principle, the latter is likely to be an easier formulation for most of us to use in day-to-day life. This principle is frequently paraphrased in different ways, including the following quote by Sir Terry Pratchett:

> 'There's no [shades of grey], only white that's got grubby. I'm surprised you don't know that. And sin, young man, is when you treat people as things. Including yourself. That's what sin is.'
>
> 'It's a lot more complicated than that -'
>
> 'No. It ain't. When people say things are a lot more complicated than that, they means they're getting worried that they won't like the truth. People as things, that's where it starts.'
>
> — Terry Pratchett, *Carpe Jugulum*

This quote nicely encapsulates both the deontological means-ends test, and the fact that deontology presents morality as a binary; acts are either moral or immoral, with no middle ground. It is also a nice demonstration of the fact that most of us have at least some familiarity with this principle, even if we may not know it by the name philosophers give it, since it is used (without naming it) in a wide range of stories and value systems. It is particularly prevalent among comic-book superheroes.

Deontology is a useful tool in large part because it makes things quite simple; because of its discrete nature and simple rules, it is easy to determine that many day-to-day actions are moral with relatively little work, after which we can proceed through our lives with confidence we are acting morally.

However, this is also one of its primary drawbacks: because there is no middle ground, it offers no clear way to weigh the relative merits of two acts which are both moral, or two acts which are both immoral. Thus, there are some types of questions to which it is not well suited to give an answer. Sometimes in life we are faced with situations where there is no truly 'good' option, and we must work to mitigate harm or choose the lesser evil; in these situations, Deontology offers us no guidance, because we are not permitted to choose any action which is not inherently good. Since "doing nothing" is also a choice, deontology can lead to a "moral bind" in which we are trapped with no permissible actions. This problem is *also* frequently used in fiction, where it is typically exploited by villains to paralyze idealistic heroes. It can occur in real life as well, though typically not in quite so dramatic a fashion.

Utilitarianism

The other framework which most people have at least some familiarity with, though possibly not by name, is *utilitarianism*, which is typically associated with John Stewart Mill. Utilitarianism is the most well known of the *consequentialist* frameworks, which state that the moral value of an act should be determined by the consequences of that act (note the contrast with deontology here). The fundamental underlying principle of utilitarianism is that 'good' can be quantified, and we should act in whatever way maximizes this quantity.

This is sometimes described as, "the greatest good for the greatest number," which has a nice ring to it but is unfortunately not well defined mathematically. What is typically used instead is "the greatest net good," or "the greatest good on average," which are mathematically equivalent for a fixed population. The word *utility* suggests that we can attempt to maximize whatever we decide has value, but in practice this framework is almost always taken as meaning that we are trying to maximize *happiness*.

The great advantage that utilitarianism presents over deontology is that it gives us a way to decide what to do for *any* situation; if we are presented with two morally good choices, we can choose the one that is the *most* good, and likewise if presented with only bad options we can choose the one which is least bad.

Unfortunately, utilitarianism also has several downsides. Firstly, doing the calculation is difficult and complex; we must determine everyone who will be impacted by a given action, then determine the net change in utility for each of them, and then take the sum of all these changes. This makes it impractical to apply to small, every-day decision making, since if we stop and do a long and involved calculation before every movement we are likely to end up just as paralyzed when faced with complex choices as we were with deontology, if for a very different reason.

Additionally, it is frequently unclear how we should go about estimating the net impact on utility to others; how do we quantify "happiness" in a way that is objective? How do we figure out the expected number of happiness-units that someone will lose if, for example, they get to the front of the line and discover we have taken the last blueberry muffin and there are none left? This requires us to know their muffin preferences, how hungry they are this morning, how accustomed they are to getting what they want, what sort of mood they are in today, etc. In general, you may be able to reason about how *you yourself* would be impacted (a line of thought which leads to the "golden rule" that we should treat others as we want to be treated), but the more dissimilar from us the other parties are, the harder it will be for us to estimate the impact on their happiness.

For these reasons, utilitarianism is often less useful in a day-to-day setting, but rather gets used for making difficult decisions with wider scale impacts that can be carefully studied and deliberated over. Even here, however, there are some potential problems, such as the fact that we are only paying attention to the *average* utility, with no consideration to its *distribution*. This can lead to decisions which utilitarianism would allow but which most people would say are not moral; for example, consider the choice between a situation in

which five people each have 10 utility (e.g. all have hard lives but are reasonably content), and a situation in which four people have 1 utility while one person has 50 utility (e.g. the first four are slaves to the last one). In the first case, the average is (5*10)/5=10, while in the second case it is (50+4)/5=10.8; the latter number is larger, so it should be preferred according to utilitarianism. However, it is difficult to argue that the latter situation is truly a morally preferable one, since most of us value fairness enough to more than offset the small increase in the average.

Another issue is that utilitarian logic relies on knowing the future impacts of our actions, but in general it is not possible to predict the future with high accuracy and confidence. This can lead to situations where we make the best decision we can with the knowledge we have at the time, but later discover that the outcome has reduced overall utility; is our decision retroactively immoral? Likewise, we must be very careful about timeframes or events that we are unable to make high precision predictions about. For example, if we consider the utility of each human life as non-zero, then the difference between eventual human extinction and the indefinite survival of the species is mathematically infinite. This means that we can justify doing literally anything if it increases our survival chance by even the smallest amount, even though most of us would feel that, for example, murdering 90% of the human population and enslaving the rest is probably not moral even if someone could convince us that doing so would lead to a 0.01% reduction in the likelihood of eventual human extinction (a line of reasoning sometimes referred to as "apocalypse morality," and again something frequently ascribed to villains in fiction).

Utilitarianism is the line of thought that leads to statements like, "the needs of the many outweigh the needs of the few," or "the ends justify the means," or "everything I do is for the greater good." As with the other frameworks, it is a useful tool, but we must be careful not to mis-use it as a rationalization to justify actions which will ultimately not benefit humanity as a whole.

Natural Justice

The *Natural Justice* ethics of John Rawles is an alternative consequentialist framework designed to address at least some of the problems with utilitarianism. In particular, it is a response to the distributional inequality that utilitarianism allows, i.e. the fact a sufficient increase in utility for a small number of individuals can outweigh a significant loss to many others.

The basic principle here is that rather than maximizing the *average* utility, we should maximize the *minimum* utility. Thus, if person A has 10 utility and person B has 5, we cannot take 3 utility away from B in order to give 4 to A (resulting in A having 14 and B having 2), but it *is* permissible to take 4 away from A to give 3 to B (resulting in A having 6 and B having 8). Note that it is also perfectly fine to do something that will increase the utility of A by 5 and increase the utility of B by 3 (resulting in A having 15 and B having 8), since this still raises the minimum.

In essence, justice ethics suggest we should prioritize the needs of those who are worst off, rather than treating everyone's utility the same way regardless of their initial state.

Rawles comes to this conclusion using a thought experiment called the "original position," in which he suggests that if all the members of a society came together to make rules without knowing what position in that society they would hold until after the fact, they would all agree on rules that emphasized fairness. After all, who would vote for a rule that benefited 1% of the population and hurt everyone else if you have a 99% chance of ending up as part of that "everyone else?"

This framework does address one of the major flaws of utilitarianism, and is a good counterpoint to it in many cases. However, there are several problems it does not address, such as the difficulty of measuring utility, the difficulty of predicting the future, etc.. It also introduces a new potential failure case, which is that we can imagine a scenario in which one grumpy old man is very resistant to having his happiness increased, leading to a situation in which everyone else bends over backwards trying to raise his happiness. For example, consider a case where nine individuals each have a utility of 10 and one has a utility of 2, and compare it to a case where nine individuals have a utility of 4 and one has a utility of 3. A strict application of natural justice would prefer the latter situation, even though most of us would feel that the dramatic sacrifice of the nine is not justified by the meager gain of the one.

Frameworks not to use

There are some "frameworks" for morality which are not useful in ethical discussions because they do not allow for meaningful argument or analysis. There are plenty of individuals who subscribe to these, but they are not valid as tools for doing formal ethical analysis. Note that this does not necessarily mean that they are entirely *wrong*, just that they aren't useful for the purpose of doing ethical analysis.

Egoism is the belief that whatever is in your personal interest is moral; that is, you believe that you are so special that one set of rules should apply to you, and another should apply to everyone else. This is not a "valid" ethical framework because there is no way to resolve disagreements; if two people who follow this framework come into conflict about what is right and what is wrong, neither will have any means to convince the other of their point of view.

Relativism is the belief that everyone's ideas about morality are equally valid, regardless of how they were arrived at (e.g., "different cultures may have different standards for morality, and that's fine"). Again, however, relativism does not offer any way to resolve a difference of opinion; if all points of view are equally valid, there can be no grounds for convincing someone who disagrees that our position has more merit.

Legalism is the belief that the morality of an act is defined fully by some form of authority, and the judgements of that authority cannot be questioned and need not be explained. This authority can take many forms, including judicial and legal systems, royal

decrees, holy books, parents, teachers, etc.. A typical argument might be something like, "this act is against the law, therefore it is wrong" (note that it makes no difference whether the law in question is secular or religious in nature). Once again, this is not a valid framework for ethical reasoning because no "reasoning" is possible; without the ability to derive the morality of an act from some agreed upon first-principle, there is no way to convince others our position is correct, nor is there any good way to determine the morality of an act which is not covered by the authority, short of asking the authority and waiting for a response (which in many cases will result in a very long wait).

Note that we may *hope* that these frameworks will produce the same conclusions as one of the valid ones above; in an ideal world, there would be perfect alignment between formally grounded ethics and the law of the land, and everyone's individual desires and moral judgements would also be aligned with this system. However, when there is a misalignment or a novel question it is important to remember that it is the *logically valid* ethical frameworks we should use when trying to determine how to resolve it.

How (not) to use ethical frameworks

As I said at the top, from my perspective the point of studying ethical frameworks is not to choose the one which is "correct," but rather to understand the strengths, weaknesses, and appropriate uses of each. Much like a carpenter would not attempt to build a table using *only* a screwdriver, when attempting to make judgements about complex problems we will typically want to use several different ethical frameworks.

For difficult and interesting problems, what is most revealing is where the different frameworks agree and where they disagree. By examining this (mis)alignment, we can help clarify our understanding of the tradeoffs involved, and thus ensure that we are making a well informed judgment that we can explain and stand behind. As an example, we may discover that the solution promoted by utilitarianism fails to pass the means-ends test, in which case we should look very carefully at why this is the case and whether the increase in net utility is worth the broader costs to society that may be hidden inside this failure.

Importantly, you should *not* simply select and use whatever framework permits whatever course of action you initially decided on. Utilitarianism is particularly vulnerable to motivated reasoning, but any ethical framework can be twisted to different ends by those who are motivated to do so; this is especially a danger to those with only a surface level familiarity with the framework in question (such as, e.g., someone who's knowledge is based primarily on this article).

It is also important to consider the stakes of the problem at hand; when making simple, every-day choices that are likely to have little broader impact, doing complex moral calculus is typically not necessary. When selecting between a blueberry muffin and a chocolate chip scone, for instance, the stakes are sufficiently low that it is not worth spending the time and mental energy on analyzing this choice using multiple ethical frameworks (unless you *really* want to annoy everyone in line behind you, in which case doing the analysis out loud is a pretty safe bet).

If the stakes are higher, however, it becomes important to do the work up front to ensure we are not causing unintentional harm or taking action we will regret in the future. For instance, deciding which of two patients should receive an organ transplant is likely to have a very significant impact on the future of those patients, and so such a decision should not be made lightly.

Likewise, deploying a change to a social media platform that will impact hundreds of millions of users is something we should think carefully about; the stakes may not be life or death for any given user, but the effects will be so widespread that even a small impact may have important consequences for the world in which we live (e.g. spreading false information about a vaccine or an election).

It is not my goal to tell you what is right and what is wrong, but rather to help you understand that there are tools you can apply to attempt to determine this for yourself. These tools will not only help you analyze the morality of decisions you might make, but can also help you understand the assumptions that underlie your sense of morality so that you are not making decisions in ignorance of the reasoning behind them.