Hello. Welcome to our second lecture on John Stuart Mill’s utilitarianism. When we finish with Mill, we’ll begin Immanuel Kant next time and we’ll finish off the course with some lectures on modern political philosophy.

Mill’s moral theory emerges directly from his conception of value. We talked last time about how our conception of value is related to our conception of obligation, that generally speaking what we take to be of the supreme value is that which we are obliged to promote and obliged to protect. Mill’s theory is a crystal clear, beautiful example of this because Mill takes happiness understood as pleasure in the absence of pain as being of supreme value. Insuring that people are happy, insuring that people experience pleasure and not pain, is our chief moral obligation.

Specifically, Mill’s moral theory says the following. I’m just gonna read it from page 7 because it’s so simple. Page 7, top of the page, quote: “The creed which accepts is the foundation of moral utility or the greatest happiness principle holds. The actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness.” That is Mill’s moral philosophy. It’s amazingly clear, simple, easy to understand, very very intuitively plausible. But underneath that simplicity, underneath that intuitive plausibility, there are some layers of complexity and difficulty that we’re gonna want to talk about.

Let’s first make sure that we understand him correctly. What Mill is saying is that what makes an action right is that it produces more happiness than unhappiness. What makes an action wrong is that it produces more unhappiness than happiness. Let’s say a few things about this. First, notice that on a view like this rightness and wrongness are matters of degree. In other words, one action can be, in a sense, more right than another. Another action can be more wrong than another. So in this view beating someone up is going to be more morally objectionable than, let’s say, swiping a piece of candy from a candy store. Because the first action causes an awful lot more
pain, an awful lot more unhappiness, than the second. Likewise, spending your life working with lepers as Mother Theresa did is going to be a greater moral act than simply giving a person who’s begging a quarter. Because the first created a lot more happiness than the second, although they both created happiness and they were both morally right as a result.

So it’s one thing to notice that on a view like Mill’s, rightness and wrongness are matters of degree. Now I take this to be a benefit. It seems to me sort of intuitively obvious that certain actions have more moral value than others. Even if both have moral value, that one can rank the moral values of actions and similarly that if actions are morally objectionable, moral objectionability is a matter of degree also, right? So certain actions are gonna be wrong — more wrong than others and this is reflected, incidentally, in the ways we punish and reward, right? I mean, if we believe that all right actions stood on equal footing and if we believe that all wrong actions stood on equal footing, we would not have differences in reward and differences in punishment. The fact that we do have differences in reward and punishment, the fact that we do reward people to different degrees depending upon the level of praiseworthiness of what they’ve done, and given that we punish people in degrees depending upon how objectionable what they’ve done is, is an indication that we intuitively take rightness and wrongness to be — praiseworthiness and objectionableness to be matters of degree and not absolute categories.

A second thing to observe is — we want to be clear about what exactly is the scope of this view. What do I mean by scope? Well, when we say that the right thing to do is that which promotes happiness, whose happiness? How many people’s happiness? For how long? Mill is not at all clear about this. However, from various places in the book where he restates this formula, it’s clear that Mill means something like the general happiness. That is, he doesn’t just mean the happiness of the
immediate actors involved in an action situation, but the more long lasting, long ranging impact of an action.

And so I think that it would be fair to make a slight addendum in order to flesh out Mill’s moral principle to say actions are right to the extent that they increase the general happiness, actions are wrong to the extent that they decrease the general happiness. Now, this still leaves the precise question of scope undetermined, but at least it makes clear that we’re just talking about the happiness of the immediate players.

The question now is how do we understand the idea of the general happiness? Let me suggest that this could mean two different things. That is, that we could read general happiness in one of two ways. The first way would be to read the idea of general happiness aggregatively. So when we say that one is morally obliged to promote the general happiness, we mean general happiness in an aggregative way. What would it mean to mean this in an aggregative way? Well, something like this. Let’s take a population of 10 people. The idea would be that the right thing to do is that which creates the most happiness amongst that entire population, however the happiness is distributed. So suppose that one action will produce — as if we can count this way. But let’s suppose that one action would produce one unit of happiness per person for a total of 10 units of happiness in the overall population. Suppose that another action only will make five people happy but it will give each of them three units of happiness for a total of 15 units in the population. On the aggregate of reading of the principle of utility, the second action is preferred to the first. Because the total amount of happiness in the population that results from the action is greater in the second case than the first, even though in the first case the happiness is more evenly distributed.

On the aggregate of reading of the principle of utility, on the aggregate of reading of Mill’s moral philosophy, what the ultimate aim of our actions should be is to promote
the general happiness in the aggregative sense, to create as much happiness in a population as is possible without respect to distribution.

Now, the aggregate of reading contrasts with what we’ll call the distributive reading. On the distributive reading what’s more important is the distribution of happiness, not the overall amount. So on the distributive reading, to go back to our example, the first action would be preferable to the second. Because even though the second action produced more happiness overall, the first action produced a broader distribution of happiness. Remember in the first example all 10 people received a unit of happiness from the action. In the second example, only five people received happiness although in greater amount. On the distributive reading, the first action is preferable because on the distributive reading of general happiness what matters is the distribution of happiness in the population, not the overall amount of happiness in the population.

There is a good deal of debate about whether Mill intended the principle of utility to be understood aggregatively or distributively. This, of course, is an historical question. That is, the question of what Mill meant by general happiness is a purely historical one. It’s of historical interest. We want to know what he thought. Notice it’s not of interest with respect to the truth. That is, which theory ultimately is the better theory depends on factors that have nothing to do with what Mill believed. It might be that Mill intended the theory to be read aggregatively but that distributive reading is a better reading.

And so the argument that’s going on — that goes on in the literature about whether Mill meant the principle of utility aggregatively or distributively is a purely historical argument. It’s an argument of historical interest.

The more interesting argument seems to be from a philosophical perspective of which is the better reading from the standpoint of moral theory. And I would say to this
that there’s no absolute answer. In response to the first question, the consensus — to
the extent that there is one. I would say it’s a slight majority. The consensus is that
Mill meant the principle of utility to be read aggregatively, not distributively. The reason
will be clear in a minute why people think this.

In terms of the question which is the better reading, which is the better way of
construing the principle of utility, I don’t know that there is an absolute answer. Both
readings have pluses and minuses. The distributive reading, of course, is the more
democratic reading. And so the distributive reading reflects a certain kind of
egalitarianism. It’s like, well, look, you know. Everybody deserves to be happy and so
we should strive to do those things that make as many people happy as possible. The
problem with the distributive reading and, of course, the problem with any sort of
majoritarian ethic is — the problem arises when what makes the majority happy makes
the minority miserable.

You see, on the aggregate of reading, even if the majority of people are happy,
so long as the misery of the smaller number of people is great enough, it will outweigh
the happiness of the majority on the aggregative reading. On the distributive reading it
will not. And so the aggregative reading of the principle utility at least in one way would
seem to be more protective of minority interests, of minority happiness. And given that
Mill in his political career was such a fervent advocate of progressive and liberal causes,
there’s good reason to think that he would’ve wanted to construe the principle of utility in
an aggregative rather than the distributive way. Indeed, in his book on liberty, much of
what he’s concerned about is the majority’s stifling of the individual, especially the
majority’s oppression of the exceptional individual. He’s very concerned on liberty with
protecting liberty against what — as the cliche goes, the tyranny of the majority.

And so we have good reason to think that Mill, at least, would’ve wanted to
construe the principle of utility in an aggregative way. Indeed, that he meant it in an
aggregative way. Because the aggregative reading of the principle of utility takes more into account minority interests, minority happiness, than does the distributive reading which is sort of ruthlessly majoritarian.

Now, one initial objection to the principle of utility to Mill’s moral theory that Mill considers is the objection that the theory is impractical. And, look, you can see this is a very sort of obvious complaint. You might say something like this. “Well, this theory is very nice, but there’s one terrible flaw with it and that is that there simply is neither enough time nor information prior to acting to be able to determine what the long-term, wide-ranging consequences of an action are gonna be.” The ordinary person in a circumstance of action simply does not have the time or the information to be able to tell in advance what are the outcomes of his action going to be. Not just on the immediate actors, not just right now, but in a broader sense over time. How can I possibly know that, the objection may go.

And so the theory is useless from a practical perspective. If this objection is vindicated, it’s a very devastating one. Because, look. The only point of having a moral theory is to apply it. The only point of having a moral theory is to derive moral rules from it that can serve as instructions for action. To simply have a moral theory so that after the fact you can either call someone moral or immoral — to have a moral theory simply to provide epithets that we fling at people is pointless and uninteresting. The reason why you want to have a moral theory is to serve as a regulator of behavior. But in order to serve that function, it’s gotta be applicable. A person has to be able to apply it in the circumstances of action. The argument here is that you can’t apply the principle of utility in the circumstances of action because the necessary information and time is simply not something that the ordinary person has in their possession.

Now, Mill’s reply, at least on the surface, it seems to me sounds pretty strong. But when we think about it a little bit, I think the reply is not adequate. Mill’s reply goes
as follows. He says, “This objection presents the moral scenarios if each person
approaches each situation fresh with no background knowledge.” But Mill wants to
say, “Look. We have all of human history, all of human experience to tell us the kinds
of results that tend to follow from certain kinds of actions.”

In other words, if I’m contemplating killing someone, it’s no mystery to me what
impact on people’s happiness killing someone will have. Surely we have all the
experience of human history to tell us that killing people tends to make people unhappy.
Obviously, the person who’s been killed is not unhappy but everybody around them is
unhappy. The people who care about them is unhappy. The people who are in the
area in which the bloodshed occurs are unhappy. And the same thing with theft and
the same thing with moral actions as well. It doesn’t take a genius to figure out that to
help someone in need is going to make them happy. Surely all of experience tells us
that.

So what Mill wants to say is look. In most cases where we’re confronted with
choices of action we don’t need to calculate the consequences. We already know what
they are. And so we have available to us what are called secondary principles or rules
of thumb. We have a rule available. We already have a rule that says don’t murder,
and the reason we have that rule is because all of human history tells us that murdering
people makes people unhappy. And so if I’m contemplating murdering somebody now,
I don’t have to figure out whether it’s gonna make people happy or unhappy. I already
know it’s gonna make people unhappy on the basis of past experience. Human
experience.

Mill voices this objection and offers this response on pages 23 to 24. Bottom of
23, quote, “Again defenders of utility often find themselves called upon to reply to such
objections as this, that there is not time previous to action for calculating and weighing
the effects of any line of conduct on the general happiness. The answer to the
objection is that there has been ample time, namely the whole past duration of the human species. During all of that time mankind has been learning by experience the tendencies of actions on which experience all the prudence as well as all the morality of life are dependent. People talk as if the commencement of this course of experience had been hitherto been put off, and as if at the moment when some man feels tempted to meddle with the property or life for another he had to begin considering for the first time whether murder and theft are injurious to human happiness.”

And, look. I think Mill certainly — in this abstract this is correct. We don’t approach a moral situation in a sense like we were just born. We approach situations knowing full well how things tend to affect people. And Mill’s allegation is we only have to calculate the happiness and unhappiness that will result from an action. In cases of actions either which have never been performed before, sort of unprecedented situations, or in the cases of action where the historical record is ambiguous — that is, it’s not clear whether more happiness or more unhappiness is created. Certainly there are many types of actions for which the human record is mixed.

Now, in either of these cases, Mill would admit, either in the case of an unprecedented situation or in the case of a situation in which the relevant actions have kind of an ambiguous history. They cause both happiness and unhappiness. Mill admits in those cases we would have to sort of calculate the happiness and unhappiness that will result from the action. But the implication is this is not a serious problem for the theory because such instances are rare.

But the problem that I have with this argument is that I — I would agree with Mill that instances of utterly unprecedented situations are rare. I think there are very few people who find themselves very often in situations that are totally unprecedented, that they have had no experience of or that no one else has had experience of that to whom they can consult. But I’m not nearly quite as sure about the second point. It seems to
me that the number of actions or types of action whose history is mixed with respect to creating happiness and unhappiness — not only is it not rare, it’s very common.

I would even say that things like killing — if you think about the last century. I mean, certainly, I guess, in an overall sense all the killing that was done in the last century caused unhappiness. But think about it in the other sense. I mean, a lot of the killing also created happiness. I am sure, absolutely certain, that the American public was very happy when the Second World War ended with the bombing of Hiroshima and Nagasaki. Very happy for the war to be over, happy to not have to engage in an invasion of the Japanese mainland, happy for many reasons. But the point is it’s not at all clear that that act created more unhappiness than happiness. Sure, it created more unhappiness for the Japanese, but I’m not sure how the utility calculations would come out if you compare that, considering how much bigger of a country the United States is than Japan and how much potentially more happiness there would be to be created in the American population.

I guess what I’m saying is I think that the argument, the objection, has a lot more bite than Mill gives it credit for because I don’t think that the number of instances in which we are confronted with actions whose history is checkered with respect to creating happiness and unhappiness -- I don’t think that that’s rare. I think that that’s common. And to the extent that it’s common means that we do have to calculate the effects of an action and then the objection that there simply is not time or information to do this prior to acting I think has great bite. So I do think to a certain extent — to a great extent there’s a certain impractibility of the theory.

But let’s consider some even more serious objections and I’m going to discuss three. One objection that I think is probably the most commonly voiced one is that utilitarianism — if one considers hypotheticals or even if one looks at human history in order to sort of make moral judgments about past events, utilitarianism often yields very
morally counterintuitive results. The theory itself I think is very intuitive. I don’t know anybody who would say that happiness is not a good — I would not know anybody who would say that a good part of our moral obligation is precisely to make people happy and not to make them unhappy. But when you actually start applying the theory to cases, both hypothetical and real, it often yields some awfully counterintuitive results, particularly in the areas with respect to distributive and retributive justice.

So let me just talk a little bit about each of these. Distributive justice speaks to everyone getting their fair share. Distributive justice is a matter of fair distribution of goods. And fair distribution of harms as well. And the example that always comes up with respect to distributive justice is sort of the lifeboat example. So imagine a ship sinking, the lifeboat only has room for a handful of people, and other people have to be thrown overboard. And imagine that being thrown overboard means you’re gonna die. I think most of us would say that distributive justice, basic fairness would require that we draw straws as to who goes. That the way of determining who gets to go in the lifeboat and who has to drown should be random. Certainly an awful lot of people would say that distributive justice requires that the process of choosing who gets to go in the lifeboat and who dies should be random.

But utilitarianism will not say that at all. What utilitarianism will want to do will be to ask, “If all the people who could drown or be saved, which ones are in the position to create the greatest amount of happiness?” So suppose that amongst the passengers there are a few extraordinary individuals. Suppose that amongst the passengers one of the people is a doctor who’s just on the brink of a cure for cancer. Suppose that another is a much beloved spiritual leader. Let’s say the head of a major church with millions and millions of congregants, loved and admired and followed by his proverbial flock. Utilitarianism would demand that we privilege those people over the ordinary Joe. So if it comes down to the doctor with the cure for cancer and some average,
ordinary, run-of-the-mill secretary for the last seat in the lifeboat, distributive justice would seem to recommend flip a coin. Utilitarianism will say chuck the secretary overboard. Because the amount of happiness that saving her life will produce is much less than the amount of happiness that saving the life of the doctor with the cure for cancer will produce. So in cases like this where issues of distributive justice are in play, utilitarianism would seem to yield morally counterintuitive results.

A similar problem arises with respect to retributive justice. Now, retributive justice means that a person should get what they deserve, their right reward and their right punishment. Hence, retributive from the word retribution. And again, in situations where retributive justice would seem to be at issue, utilitarianism often yields morally counterintuitive results. So let’s continue on with a similar case with some similar characters.

We have a man, a famous doctor, who is on the brink of a cure for cancer. Suppose that this man murders his wife. He can’t stand her, he wants to get rid of her, wants the insurance money, etc., etc., etc. He murders his wife and imagine that he’s caught by the police. And suppose that his guilt is beyond any reasonable doubt. He’s proven guilty. His hands are all over it. Fingerprints are all over everything. Retributive justice would seem to obviously require that the man be punished with whatever is the appropriate punishment. But suppose that the man says the following. He says, “Look. I have the cure for cancer. You know this. If you do anything to me — if you imprison me, if you so much as fine me — I will never, ever, ever tell you the secret to curing cancer. I have it written down. It’s locked away somewhere. You’re never gonna find it. You’re never gonna get in and find the cure. Only I know what it is. I’ve never told anyone else. If you do anything to punish me, I will forever keep my silence.”

Now, retributive justice would seem to require, “Well, too bad for everyone, but
you murdered somebody and you’ve gotta pay the price. That’s your just punishment.” But utilitarianism would say that we’re morally obliged to free the man. Now, if you don’t find this outrageous, considering a person has committed exactly the same crime but has nothing valuable to impart, no valuable information to impart, he goes straight to the electric chair. So there actually also here arise issues of distributive justice equals sort of distribution of punishment as well as rewards. Well, the guy might say, “Well, wait a minute. Why do I have to fry in the electric chair when this guy gets to go free?” Well, this would seem obviously unjust. And yet the utilitarian would say — would have to say, “It’s morally obligatory to free the doctor and, on the other hand, fry the other guy.” And so again, in instances where retributive justice is at issue, utilitarianism will often yield morally counterintuitive results.

This is one form of objection. There are so many ways of attempting to deal with this objection that we can’t list them all. I’m just sort of gonna register these objections. I’m not going to make much of a point about which ones I think succeed or fail or what all the replied have been because they’re too numerous to go into in an introductory level course.

Let me sort of mention another objection that arises. This is an objection — actually, this and the last objection are objections that are going to anticipate some of the things that Kant is going to say. In a lot of ways Mill’s moral philosophy and Kant’s are complementary. Each one addresses morally — sort of significant characteristics of morality that the other one misses. If you could somehow put them together, you’d have the best moral theory in the world. The problem is that the theories are, to a certain extent, contradictory. But each theory seems to thrive where the other one withers. And these last two objections to utilitarianism are points on which Kantianism is a much stronger moral philosophy.

So let’s take the first one. The utilitarian is focused entirely on the outcomes of
your actions. Utilitarianism is, as we’ve just defined it, a consequentialist moral philosophy. Consequentialism is simply the view that the moral value of an act lies outside the act specifically in the action’s consequences. For the utilitarian, what makes an action right or wrong is whether it produces more happiness or unhappiness. The consequences of the action are what determine its moral value.

Now, certainly part of what’s going to determine whether the consequences of one’s actions are more happiness or unhappiness, part of what’s going to determine that are what you intend to have happen. But notice, of course, that circumstances, chance, luck are always gonna play a role and often a very large role. I guess what I’m trying to say is it’s not at all clear just how much control you have over the outcomes of your action in any given circumstance. Sure, you can control what you want to have happen. You can control what you do. But what actual effect it has is subject — is a matter of not only what you’ve intended and what you’ve done but of the surrounding circumstances over which you may have no control whatsoever.

I’m sure you’ve heard of the expression “unintended consequences.” Well, unintended consequences are legion. The point just is that it would seem that the outcomes of our actions, whether an action causes happiness or unhappiness, is to a large extent hostage to fortune. Not within our control. We can control what we intend, we can control what we do. We have much less control over what are the outcomes of what we do.

And this would seem to play — and, now, remember. The moral value of an action from Mill is based entirely on the outcome. Which means that the moral value of your action is based largely on things that are out of your control. And so you are being blamed or praised for things that are largely a result of factors over which you have no control. This is morally counterintuitive. It seems to me, at least, that a person should only be blamed or praised for those things over which they had control.
To blame or praise someone over something that they had no control over would seem to just sort of be wrong, right?

And this is a serious problem. This is sort of the problem with moral luck — the problem of moral luck. The utilitarian ties moral value to outcomes in such a way that the moral judgment of a person and their actions is going to be based largely on factors that are out of their control, that they in a sense had nothing directly to do with. Many people have found this to be highly objectionable, objectionable because it clashes with our ordinary intuitions about how praise and blame are to be ascribed. We typically don’t praise people for things they didn’t do on purpose. If somebody — look. We praise somebody who, let’s say, through brilliant effort and talent and ingenuity and effort achieves success. We don’t praise someone who wins the lottery. I mean, we might be happy for them but they won by accident. Something good happened by accident. We tend not to blame people for things that are not the results of their own volition.

And so the fact that the most important factor in the moral evaluation of an action is something that rests on circumstances that are largely out of an agent’s control subjects moral value to a kind of fortune or luck which is thought by many to be counterintuitive and problematic.

And while we’re on the subject of intentions, let’s just sort of make the last point. Again, this will point us directly towards Kant’s ethics. On the utilitarian view, intentions are absolutely irrelevant to the moral value of one’s action. This again would seem to just be crazy, right? Surely your intentions are relevant to the moral assessment of what you’ve done.

Think, for example — let’s take another case. Imagine that I’m walking down the street and I see a woman drowning in a pond. And suppose I rushed to save the woman and I pull her out of the pond and I save her life. And so certainly the
consequence of my action is a good one. My action created more happiness than unhappiness. But let’s suppose that I did this for entirely selfish reasons. Suppose I actually don’t care whether the woman lives or dies. The reason I saved her was that I saw an opportunity to become a local celebrity. I thought hmmm, you know. If I save that woman from drowning, I’ll be in the newspapers. I’ll be on TV. I’ll be a local celebrity. People will love me. I’ll get all kinds of kudos. I’ll get gifts in the mail. The major will give me the key to the city, and so on and so forth. So suppose that I did something that has a positive outcome but that I did it for utterly craven, selfish reasons.

Now imagine another person. Exact same situation. The person walks by a pond, sees a woman drowning, saves the woman’s life. In other words, suppose that the same person’s action has exactly the same outcome as mine, as the first. But suppose that this second person saved the woman because he has a deep-seated love for human life. He believes that that woman has the right to live and he’s gonna do everything in his power, even risk his own life, to save her.

Now, according to utilitarianism, our two actions have identical moral values. The only thing relevant to the assessment of the moral value of an action is the amount of happiness that it produces. But I think that every one of us would want to say that the second person’s action was more morally worthy than the first because of the intentions on which the second person acted opposed to the first. And what this would suggest is that intention is certainly morally relevant to the value of an action and to the assessment of a person.

This is sort of one of utilitarianism’s great acknowledged weaknesses. It leaves no room for the moral relevance of intention. Now, Kant’s moral philosophy is going to leave room for that. The problem with Kant’s moral philosophy is that it leaves no room in the moral equation for outcomes, for consequences. In a sense, Kant and Mill’s problems are mirror images of each other. Mill has what Kant lacks and Kant has what
Mill lacks. This strikes me as sort of an interesting situation and it suggests that they each have a part of the truth about right and wrong, albeit not the whole of it.

I want you to start reading Kant and Kant is very difficult, much more difficult to read than Mill. Mill, of course, is writing in English and is a clear, straightforward writer. Kant writes in German so this is in translation, and even in German is known to be a very convoluted, complicated, turgid sort of writer. There's a few things I just want you to pay attention to. I first want you to see if you can explain — earlier on in the course we described the difference between a consequentialist's moral theories, on the one hand, and deontological moral theories on the other. I want you to try and explain the sense in which Kant — what makes Kant’s moral philosophy deontological rather than consequentialist.

Second, I want you to pay attention to the role that intentions and principles play in Kant’s moral philosophy. Also in the role that reason plays. See if you can, in let's say a short paragraph, explain in what sense does Kant think morality is rational and what sense is our moral duty a requirement of reason or an obligation of reason. These are all things I want you to think about for next time.

So we will begin with Kant next time and I'll see you then.