

from Fred Feldman,
Introductory Ethics
(Prentice-Hall, 1978).

7

Kant I

Sometimes our moral thinking takes a decidedly nonutilitarian turn. That is, we often seem to appeal to a principle that is inconsistent with the whole utilitarian standpoint. One case in which this occurs clearly enough is the familiar tax-cheat case. A person decides to cheat on his income tax, rationalizing his misbehavior as follows: "The government will not be injured by the absence of my tax money. After all, compared with the enormous total they take in, my share is really a negligible sum. On the other hand, I will be happier if I have the use of the money. Hence, no one will be injured by my cheating, and one person will be better off. Thus, it is better for me to cheat than it is for me to pay."

In response to this sort of reasoning, we may be inclined to say something like this: "Perhaps you are right in thinking that you will be better off if you cheat. And perhaps you are right in thinking that the government won't even know the difference. Nevertheless, your act would be wrong. For if everyone were to cheat on his income taxes, the government would soon go broke. Surely you can see that you wouldn't want others to act in the way you propose to act. So you shouldn't act in that way." While it may not be clear that this sort of response would be decisive, it should be clear that this is an example of a sort of response that is often given.

There are several things to notice about this response. For one, it is not based on the view that the example of the tax cheat will provoke everyone else to cheat too. If that were the point of the response, then the response might be explained on the basis of utilitarian considerations. We could understand the responder to be saying that the tax cheater has miscalculated his utilities. Whereas he thinks his act of cheat-

ing has high utility, in fact it has low utility because it will eventually result in the collapse of the government. It is important to recognize that the response presented above is not based upon any such utilitarian considerations. This can be seen by reflecting on the fact that the point could just as easily have been made in this way: "Of course, very few other people will know about your cheating, and so your behavior will not constitute an example to others. Thus, it will not provoke others to cheat. Nevertheless, your act is wrong. For if everyone were to cheat as you propose to do, then the government would collapse. Since you wouldn't want others to behave in the way you propose to behave, you should not behave in that way. It would be wrong to cheat."

Another thing to notice about the response in this case is that the responder has not simply said, "What you propose to do would be cheating; hence, it is wrong." The principle in question is not simply the principle that cheating is wrong. Rather, the responder has appealed to a much more general principle, which seems to be something like this: If you wouldn't want everyone else to act in a certain way, then you shouldn't act in that way yourself.

This sort of general principle is in fact used quite widely in our moral reasoning. If someone proposes to remove the pollution-control devices from his automobile, his friends are sure to say "What if everyone did that?" They would have in mind some dire consequences for the quality of the air, but their point would not be that the removal of the pollution-control device by one person will in fact cause others to remove theirs, and will thus eventually lead to the destruction of the environment. Their point, rather, is that if their friend would not want others to act in the way he proposes to act, then it would be wrong for him to act in that way. This principle is also used against the person who refrains from giving to charity; the person who evades the draft in time of national emergency; the person who tells a lie in order to get out of a bad spot; and even the person who walks across a patch of newly seeded grass. In all such cases, we feel that the person acts wrongly not because his actions will have bad results, but because he wouldn't want others to behave in the way he behaves.

A highly refined version of this nonutilitarian principle is the heart of the moral theory of Immanuel Kant.¹ In his *Groundwork of the Metaphysic of Morals*,² Kant presents, develops, and defends the thesis that something like this principle is the "supreme principle of morality." Kant's presentation is rather complex; in parts, it is very hard to follow. Part of the trouble arises from his use of a rather unfamiliar technical vocabulary. Another source of trouble is that Kant is concerned with establishing a variety of other points in this little book, and some of these involve fairly complex issues in metaphysics and epistemology. Since our aim here is simply to present a clear, concise account of Kant's basic

moral doctrine, we will have to ignore quite a bit of what he says in the book.

Kant formulates his main principle in a variety of different ways. All of the members of the following set of formulations seem to have a lot in common:

I ought never to act except in such a way that I can also will that my maxim should become a universal law.³

Act only on that maxim through which you can at the same time will that it should become a universal law.⁴

Act as if the maxim of your action were to become through your will a universal law of nature.⁵

We must be able to will that a maxim of our action should become a universal law—this is the general canon for all moral judgment of action.⁶

Before we can evaluate this principle, which Kant calls the *categorical imperative*, we have to devote some attention to figuring out what it is supposed to mean. To do this, we must answer a variety of questions. What is a maxim? What is meant by "universal law"? What does Kant mean by "will"? Let us consider these questions in turn.

Maxims

In a footnote, Kant defines *maxim* as "a subjective principle of volition."⁷ This definition is hardly helpful. Perhaps we can do better. First, however, a little background.

Kant apparently believes that when a person engages in genuine action, he always acts on some sort of general principle. The general principle will explain what the person takes himself to be doing and the circumstances in which he takes himself to be doing it. For example, if I need money, and can get some only by borrowing it, even though I know I won't be able to repay it, I might proceed to borrow some from a friend. My maxim in performing this act might be, "Whenever I need money and can get it by borrowing it, then I will borrow it, even if I know I won't be able to repay it."

Notice that this maxim is *general*. If I adopt it, I commit myself to behaving in the described way *whenever* I need money and the other conditions are satisfied. In this respect, the maxim serves to formulate a general principle of action rather than just some narrow reason applicable in just one case.⁸ So a maxim must describe some general sort of situation, and then propose some form of action for the situation. To adopt a maxim is to commit yourself to acting in the described way whenever the situation in question arises.

It seems clear that Kant holds that every action has a maxim, although he does not explicitly state this view. When we speak of an action

here, we mean a concrete, particular action, or *act-token*, rather than an *act-type*. Furthermore, we must distinguish between genuine actions and what we may call "mere bodily movements." It would be absurd to maintain that a man who scratches himself in his sleep is acting on the maxim "When I itch, I shall scratch." His scratching is a mere bodily movement, and has no maxim. A man who deliberately sets out to borrow some money from a friend, on the other hand, does perform an action. And according to our interpretation of Kant, his action must have a maxim.

It would be implausible to maintain that before we act, we always consciously formulate the maxim of our action. Most of the time we simply go ahead and perform the action without giving any conscious thought to what we're doing, or what our situation is. We're usually too intent on getting the job done. Nevertheless, if we are asked after the fact, we often recognize that we actually were acting on a general policy, or maxim. For example, if you are taking a test, and you set about to answer each question correctly, you probably won't give any conscious thought to your maxim. You will be too busy thinking about the test. But if someone were to ask you to explain what you are doing and to explain the policy upon which you are doing it, you might then realize that in fact you have been acting a maxim. Your maxim might be, "Whenever I am taking an academic test, and I believe I know the correct answers, I shall give what I take to be the correct answers." So a person may act on a maxim even though she hasn't consciously entertained it.

In one respect, the maxim of an action may be inaccurate: it does not so much represent the actual situation of the action as it does the situation the agent takes himself to be in. Suppose, for example, that I have a lot of money in my savings account but I have forgotten all about it. I take myself to be broke. When I go out to borrow some money from a friend, my maxim might be, "When I am broke and can get money in no other way, I shall borrow some from a friend." In this case, my maxim does not apply to my actual situation. For my actual situation is not one in which I am broke. Yet the maxim does apply to the situation I take myself to be in. For I believe that I am broke, and I believe that I can get money in no other way. So it is important to recognize that a maxim is a general policy statement that describes the sort of situation the agent takes himself to be in when he performs an action, and the sort of action he takes himself to be performing. In fact, both the situation and the action may be different from what the agent takes them to be.

Another point about maxims that should be recognized is this. Externally similar actions may in fact have radically different maxims. Here is an elaborated version of an example given by Kant that illustrates this point.⁹ Suppose there are two grocers, Mr. Grimbey and Mr. Hughes. Mr. Grimbey's main goal in life is to get rich. After careful consideration, he has decided that in the long run he'll make more money if he

gains a reputation for treating his customers fairly. In other words, he believes that "honesty is the best policy—because it pays." Hence, Mr. Grimbey scrupulously sees to it that every customer gets the correct change. When Mr. Grimbey gives correct change to a customer, he acts on this maxim:

*M*₁: When I can gain a good business reputation by giving correct change, I shall give correct change.

Mr. Hughes, on the other hand, has decided that it would be morally wrong to cheat his customers. This decision has moved him to adopt the policy of always giving the correct change. He doesn't care whether his honest dealings will in the long run contribute to an increase in sales. Even if he were to discover that honesty in business dealings does *not* pay, he would still treat his customers honestly. So Mr. Hughes apparently acts on some maxim such as this:

*M*₂: When I can perform a morally right act by giving correct change, I shall give correct change.

Mr. Grimbey's overt act of giving correct change to a customer looks just like Mr. Hughes's overt act of giving correct change to a customer. Their customers cannot tell, no matter how closely they observe the behavior of Mr. Grimbey and Mr. Hughes, what their maxims are. However, as we have seen, the actions of Mr. Grimbey are associated with a maxim radically different from that associated with the actions of Mr. Hughes.

For our purposes, it will be useful to introduce a concept that Kant does not employ. This is the concept of the *generalized form* of a maxim. Suppose I decide to go to sleep one night and my maxim in performing this act is this:

*M*₃: Whenever I am tired, I shall sleep.

My maxim is stated in such a way as to contain explicit references to me. It contains two occurrences of the word "I." The generalized form of my maxim is the principle we would get if we were to revise my maxim so as to make it applicable to everyone. Thus, the generalized form of my maxim is this:

*GM*₃: Whenever anyone is tired, he will sleep.

In general, then, we can represent the form of a maxim in this way:

M: Whenever I am ———, I shall ———.

Actual maxims have descriptions of situations in the first blank and descriptions of actions in the second blank. The generalized form of a maxim can be represented in this way:

GM: Whenever anyone is ———, she will ———.

So much, then, for maxims. Let us turn to our second question, "What is meant by universal law?"

Universal Law

When, in the formulation of the categorical imperative, Kant speaks of "universal law," he seems to have one or the other of two things in mind. Sometimes he seems to be thinking of a *universal law of nature*, and sometimes he seems to be thinking of a *universal law of freedom*.

A *law of nature* is a fully general statement that describes not only how things are, but how things always *must* be. Consider this example: If the temperature of a gas in an enclosed container is increased, then the pressure will increase too. This statement accurately describes the behavior of gases in enclosed containers. Beyond this, however, it describes behavior that is, in a certain sense, necessary. The pressure not only *does* increase, but it *must* increase if the volume remains the same and the temperature is increased. This "must" expresses not logical or moral necessity, but "physical necessity." Thus, a law of nature is a fully general statement that expresses a physical necessity.

A *universal law of freedom* is a universal principle describing how all people ought to act in a certain circumstance. It does not have to be a legal enactment—it needn't be passed by Congress or signed by the president. Furthermore, some universal laws of freedom are not always followed—although they should be. If in fact it is true that all promises ought to be kept, then this principle is a universal law of freedom: If anyone has made a promise, he keeps it. The "must" in a statement such as "If you have made a promise, then you must keep it" does not express logical or physical necessity. It may be said to express moral necessity. Using this concept of moral necessity, we can say that a universal law of freedom is a fully general statement that expresses a moral necessity.

Sometimes Kant's categorical imperative is stated in terms of universal laws of nature, and sometimes in terms of universal laws of freedom. We will consider the "law of nature" version, since Kant appeals to it in discussing some fairly important examples.

Willing

To will that something be the case is more than to merely wish for it to be the case. A person might wish that there would be peace everywhere in the world. Yet knowing that it is not within his power to bring about this wished-for state of affairs, he might refrain from willing that there be

peace everywhere in the world. It is not easy to say just what a person does when he wills that something be the case. According to one view, willing that something be the case is something like commanding yourself to make it be the case. So if I will my arm to go up, that would be something like commanding myself to raise my arm. The Kantian concept of willing is a bit more complicated, however. According to Kant, it makes sense to speak of willing something to happen, even if that something is not an action. For example, we can speak of someone willing that everyone keep their promises.

Some states of affairs are impossible. They simply cannot occur. For example, consider the state of affairs of your jumping up and down while remaining perfectly motionless. It simply cannot be done. Yet a sufficiently foolish or irrational person might will that such a state of affairs occur. That would be as absurd as commanding someone else to jump up and down while remaining motionless. Kant would say of a person who has willed in this way that his will has "contradicted itself." We can also put the point by saying that the person has willed inconsistently.

Inconsistency in willing can arise in another, somewhat less obvious way. Suppose a person has already willed that he remain motionless. He does not change this volition, but persists in willing that he remain motionless. At the same time, however, he begins to will that he jump up and down. Although each volition is self-consistent, it is inconsistent to will both of them at the same time. This is a second way in which inconsistency in willing can arise.

It may be the case that there are certain things that everyone must always will. For example, we may have to will that we avoid intense pain. Anyone who wills something that is inconsistent with something everyone must will, thereby wills inconsistently.

Some of Kant's examples suggest that he held that inconsistency in willing can arise in a third way. This form of inconsistency is a bit more complex to describe. Suppose a person wills to be in Boston on Monday and also wills to be in San Francisco on Tuesday. Suppose, furthermore, that because of certain foul-ups at the airport it will be impossible for her to get from Boston to San Francisco on Tuesday. In this case, Kant would perhaps say that the person has willed inconsistently.

In general, we can say that a person wills inconsistently if he wills that p be the case and he wills that q be the case and it is impossible for p and q to be the case together.

The Categorical Imperative

With all this as background, we may be in a position to interpret the first version of Kant's categorical imperative. Our interpretation is this:

*CI*₁: An act is morally right if and only if the agent of the act can consistently will that the generalized form of the maxim of the act be a law of nature.

We can simplify our formulation slightly by introducing a widely used technical term. We can say that a maxim is *universalizable* if and only if the agent who acts upon it can consistently will that its generalized form be a law of nature. Making use of this new term, we can restate our first version of the categorical imperative as follows:

CI'₁: An act is morally right if and only if its maxim is universalizable.

As formulated here, the categorical imperative is a statement of necessary and sufficient conditions for the moral rightness of actions. Some commentators have claimed that Kant did not intend his principle to be understood in this way. They have suggested that Kant meant it to be understood merely as a necessary but not sufficient condition for morally right action. Thus, they would prefer to formulate the imperative in some way such as this:

CI''₁: An act is morally right only if its maxim is universalizable.

Understood in this way, the categorical imperative points out one thing to avoid in action. That is, it tells us to avoid actions whose maxims cannot be universalized. But it does not tell us the distinguishing feature of the actions we should perform. Thus, it does not provide us with a criterion of morally right action. Since Kant explicitly affirms that his principle is "the supreme principle of morality," it is reasonable to suppose that he intended it to be taken as a statement of necessary and sufficient conditions for morally right action. In any case, we will take the first version of the categorical imperative to be *CI*₁, rather than *CI*''₁.

It is interesting to note that other commentators have claimed that the categorical imperative isn't a criterion of right action at all. They have claimed that it was intended to be understood as a criterion of correctness for *maxims*.¹⁰ These commentators might formulate the principle in this way:

CI'''₁: A maxim is morally acceptable if and only if it is universalizable.

This interpretation is open to a variety of objections. In the first place, it is not supported by the text. Kant repeatedly states that the categorical imperative is the basic principle by which we are to evaluate actions.¹¹ Furthermore, when he presents his formulations of the

categorical imperative, he generally states it as a principle about the moral rightness of action. Finally, it is somewhat hard to see why we should be interested in a principle such as *CI*'''₁. For it does not constitute a theory about right action, or good persons, or anything else that has traditionally been a subject of moral enquiry. *CI*₁, on the other hand, competes directly with act utilitarianism, rule utilitarianism, and other classical moral theories.

In order to gain a better insight into the workings of the categorical imperative, it may be worthwhile to compare it with a doctrine with which it is sometimes confused—the golden rule. The golden rule has been formulated in a wide variety of ways.¹² Generally, however, it looks something like this:

GR: An act is morally right if and only if, in performing it, the agent refrains from treating others in ways in which he would not want the others to treat him.

According to *GR*, then, if you wouldn't want others to lie to you, it is wrong to lie to them. If you would want others to treat you with respect, then it is right to treat others with respect.

Kant explicitly rejects the view that his categorical imperative is equivalent to the golden rule.¹³ He points out a number of respects in which the two doctrines differ. For one, *GR* is not applicable to cases in which only one person is involved. Consider suicide. When a person commits suicide, he does not "treat others" in any way; he only "treats himself." Hence, when a person commits suicide, he does not treat others in ways in which he would not want the others to treat him. Therefore, under *GR*, anyone who commits suicide performs a morally right act. *CI*₁, on the other hand, may not yield this result. For if a person commits suicide, he does so on a maxim, whether other people are involved or not. Either his maxim is universalizable, or it is not. If it is not, *CI*₁ entails that his action is not right. If it is, *CI*₁ entails that his action is right. In this respect, *CI*₁ is clearly distinct from *GR*.

Kant also hints at another respect in which the two doctrines differ. Suppose a person considers herself to be utterly self-sufficient. She feels that she has no need of aid from others. *GR* then has nothing to say against her refraining from extending any kindness to others. After all, she has no objection to being treated in this unkind way by them. So *GR* entails that her behavior is morally right. *CI*₁, on the other hand, has no such consequence. Whether this person is willing to be mistreated by others or not, it may still be irrational of her to will that it be a law of nature that no one help anyone else. If so, *CI*₁ rules out uncharitableness, whether the agent likes it or not.

Similar considerations apply to masochists, whose behavior is not

adequately guided by GR. After all, we surely don't want to allow the masochist to torture others simply on the grounds that he wouldn't object to being tortured by them! The unusual desires of masochists do not pose any special threat to CI₁.

So the main difference between GR and CI₁ seems to be this: According to GR, what makes an act right is the fact that the agent would not object to "having it done to himself." This opens the door to incorrect results in cases in which the agent, for some unexpected reason, would not object to being mistreated. According to CI₁, what makes an act right is the fact that the agent's maxim in performing it can be universalized. Thus, even if he would not object to being mistreated by others, his mistreatment of them may be wrong simply because it would be *irrational* to will that everyone should mistreat others in the same way.

Kant's Four Examples

In a very famous passage in Chapter II of the *Groundwork*, Kant presents four illustrations of the application of the categorical imperative.¹⁴ In each case, in Kant's opinion, the act is morally wrong and the maxim is not universalizable. Thus, Kant holds that his theory implies that each of these acts is wrong. If Kant is right about this, then he has given us four positive instances of his theory. That is, he has given us four cases in which his theory yields correct results. Unfortunately, the illustrations are not entirely persuasive.

Kant distinguishes between "duties to self" and "duties to others." He also distinguishes between "perfect" and "imperfect" duties. This gives him four categories of duty: "perfect to self," "perfect to others," "imperfect to self," and "imperfect to others." Kant gives one example of each type of duty. By "perfect duty," Kant says he means a duty "which admits of no exception in the interests of inclination."¹⁵ Kant seems to have in mind something like this: If a person has a perfect duty to perform a certain kind of action, then he must *always* do that kind of action when the opportunity arises. For example, Kant apparently holds that we must always perform the (negative) action of refraining from committing suicide. This would be a perfect duty. On the other hand, if a person has an imperfect duty to do a kind of action, then he must at least *sometimes* perform an action of that kind when the opportunity arises. For example, Kant maintains that we have an imperfect duty to help others in distress. We should devote at least some of our time to charitable activities, but we are under no obligation to give all of our time to such work.

The perfect/imperfect distinction has been drawn in a variety of ways—none of them entirely clear. Some commentators have said that if

a person has a perfect duty to do a certain action, *a*, then there must be someone else who has a corresponding right to demand that *a* be done. This seems to be the case in Kant's second example, but not in his first example. Thus, it isn't clear that we should understand the concept of perfect duty in this way. Although the perfect/imperfect distinction is fairly interesting in itself, it does not play a major role in Kant's theory. Kant introduces the distinction primarily to insure that his examples will illustrate different kinds of duty.

Kant's first example illustrates the application of CI₁ to a case of perfect duty to oneself—the alleged duty to refrain from committing suicide. Kant describes the miserable state of the person contemplating suicide, and tries to show that his categorical imperative entails that the person should not take his own life. In order to simplify our discussion, let us use the abbreviation "a₁" to refer to the act of suicide the man would commit, if he were to commit suicide. According to Kant, every act must have a maxim. Kant tells us the maxim of a₁: "From self-love I make it my principle to shorten my life if its continuance threatens more evil than it promises pleasure."¹⁶ Let us simplify and clarify this maxim, understanding it as follows:

M(a₁): When continuing to live will bring me more pain than pleasure, I shall commit suicide out of self-love.

The generalized form of this maxim is as follows:

GM(a₁): Whenever continuing to live will bring anyone more pain than pleasure, he will commit suicide out of self-love.

Since Kant believes that suicide is wrong, he attempts to show that his moral principle, the categorical imperative, entails that a₁ is wrong. To do this, of course, he needs to show that the agent of a₁ cannot consistently will that *GM*(a₁) be a law of nature. Kant tries to show this in the following passage:

... a system of nature by whose law the very same feeling whose function is to stimulate the furtherance of life should actually destroy life would contradict itself and consequently could not subsist as a system of nature. Hence this maxim cannot possibly hold as a universal law of nature and is therefore entirely opposed to the supreme principle of all duty.¹⁷

The general outline of Kant's argument is clear enough:

Suicide Example

- (1) *GM*(a₁) cannot be a law of nature.
- (2) If *GM*(a₁) cannot be a law of nature, then the agent of a₁ cannot consistently will that *GM*(a₁) be a law of nature.

(3) a_1 is morally right if and only if the agent of a_1 can consistently will that $GM(a_1)$ be a law of nature.

(4) Therefore, a_1 is not morally right.

In order to determine whether Kant really has shown that his theory entails that a_1 is not right, let us look at this argument more closely. First of all, for our purposes we can agree that the argument is valid. If all the premises are true, then the argument shows that the imagined act of suicide would not be right. CI_1 , here being used as premise (3), would thus be shown to imply that a_1 is not right.

Since we are now interested primarily in seeing how Kant makes use of CI_1 , we can withhold judgment on the merits of it for the time being.

The second premise seems fairly plausible. For although an irrational person could probably will almost anything, it surely would be difficult for a perfectly rational person to will that something be a law of nature if that thing could not be a law of nature. Let us grant, then, that it would not be possible for the agent to consistently will that $GM(a_1)$ be a law of nature if in fact $GM(a_1)$ could not be a law of nature.

The first premise is the most troublesome. Kant apparently assumes that "self-love" has as its function, the stimulation of the furtherance of life. Given this, he seems to reason that self-love cannot also contribute sometimes to the destruction of life. Perhaps Kant assumes that a given feeling cannot have two "opposite" functions. However, if $GM(a_1)$ were a law of nature, self-love would have to contribute toward self-destruction in some cases. Hence, Kant seems to conclude, $GM(a_1)$ cannot be a law of nature. And so we have our first premise.

If this is Kant's reasoning, it is not very impressive. In the first place, it is not clear why we should suppose that self-love has the function of stimulating the furtherance of life. Indeed, it is not clear why we should suppose that self-love has any function at all! Second, it is hard to see why self-love can't serve two "opposite" functions. Perhaps self-love motivates us to stay alive when continued life would be pleasant, but motivates us to stop living when continued life would be unpleasant. Why should we hold this to be impossible?

So it appears that Kant's first illustration is not entirely successful. Before we turn to the second illustration, however, a few further comments may be in order. First, some philosophers would say that it is better that Kant's argument failed here. Many moralists would take the following position: Kant's view about suicide is wrong. The act of suicide out of self-love, a_1 , is morally blameless. In certain circumstances suicide is each person's "own business." Thus, these moralists would say that if the categorical imperative did imply that a_1 is morally wrong, as Kant tries to show, then Kant's theory would be defective. But since Kant was

not entirely successful in showing that his theory had this implication, the theory has not been shown to have any incorrect results.

A second point to notice about the suicide example is its scope. It is important to recognize that in this passage Kant has not attempted to show that suicide is always wrong. Perhaps Kant's personal view is that it is never right to commit suicide. However, in the passage in question he attempts to show only that a certain act of suicide, one based on a certain maxim, would be wrong. For all Kant has said here, other acts of suicide, done according to other maxims, might be permitted by the categorical imperative.

Let us turn now to the second illustration. Suppose I find myself hard-pressed financially and I decide that the only way in which I can get some money is by borrowing it from a friend. I realize that I will have to promise to repay the money, even though I won't in fact be able to do so. For I foresee that my financial situation will be even worse later on than it is at present. If I perform this action, a_2 , of borrowing money on a false promise, I will perform it on this maxim:

$M(a_2)$: When I need money and can get some by borrowing it on a false promise, then I shall borrow the money and promise to repay, even though I know that I won't be able to repay.

The generalized form of my maxim is this:

$GM(a_2)$: Whenever anyone needs money and can get some by borrowing it on a false promise, then he will borrow the money and promise to repay, even though he knows that he won't be able to repay.

Kant's view is that I cannot consistently will that $GM(a_2)$ be a law of nature. This view emerges clearly in the following passage:

... I can by no means will a universal law of lying; for by such a law there could properly be no promises at all, since it would be futile to profess a will for future action to others who would not believe my profession or who, if they did so over-hastily, would pay me back in like coin; and consequently my maxim, as soon as it was made a universal law, would be bound to annul itself.¹⁸

It is important to be clear about what Kant is saying here. He is not arguing against lying on the grounds that if I lie, others will soon lose confidence in me and eventually won't believe my promises. Nor is he arguing against lying on the grounds that my lie will contribute to a general practice of lying, which in turn will lead to a breakdown of trust and the destruction of the practice of promising. These considerations

are basically utilitarian. Kant's point is more subtle. He is saying that there is something covertly self-contradictory about the state of affairs in which, as a law of nature, everyone makes a false promise when in need of a loan. Perhaps Kant's point is this: Such a state of affairs is self-contradictory because, on the one hand, in such a state of affairs everyone in need would borrow money on a false promise, and yet, on the other hand, in that state of affairs no one could borrow money on a false promise—for if promises were always violated, who would be silly enough to loan any money?

Since the state of affairs in which everyone in need borrows money on a false promise is covertly self-contradictory, it is irrational to will it to occur. No one can consistently will that this state of affairs should occur. But for me to will that $GM(a_2)$ be a law of nature is just for me to will that this impossible state of affairs occur. Hence, I cannot consistently will that the generalized form of my maxim be a law of nature. According to CI_1 , my act is not right unless I can consistently will that the generalized form of its maxim be a law of nature. Hence, according to CI_1 , my act of borrowing the money on the false promise is not morally right.

We can restate the essentials of this argument much more succinctly:

Lying-promise example

- (1) $GM(a_2)$ cannot be a law of nature.
- (2) If $GM(a_2)$ cannot be a law of nature, then I cannot consistently will that $GM(a_2)$ be a law of nature.
- (3) a_2 is morally right if and only if I can consistently will that $GM(a_2)$ be a law of nature.
- (4) Therefore, a_2 is not morally right.

The first premise is based upon the view that it would somehow be self-contradictory for it to be a law of nature that everyone in need makes a lying promise. For in that (allegedly impossible) state of affairs there would be promises, since those in need would make them, and there would also not be promises, since no one would believe that anyone was really committing himself to future payment by the use of the words "I promise." So, as Kant says, the generalized form of the maxim "annuls itself." It cannot be a law of nature.

The second premise is just like the second premise in the previous example. It is based on the idea that it is somehow irrational to will that something be the case if in fact it is impossible for it to be the case. So if it really is impossible for $GM(a_2)$ to be a law of nature, then it would be irrational of me to will that it be so. Hence, I cannot consistently will that the generalized form of my maxim be a law of nature. In other words, I

cannot consistently will that it be a law of nature that whenever anyone needs money and can get some on a false promise, then he will borrow some and promise to repay, even though he knows that he won't be able to repay.

The third premise of the argument is the categorical imperative. If the rest of the argument is acceptable, then the argument as a whole shows that the categorical imperative, together with these other facts, implies that my lying promise would not be morally right. This would seem to be a reasonable result.

Some readers have apparently taken this example to show that according to Kantianism, it is always wrong to make a false promise. Indeed, Kant himself may have come to this conclusion. Yet if we reflect on the argument for a moment, we will see that the view of these readers is surely not the case. At best, the argument shows only that one specific act of making a false promise would be wrong. That one act is judged to be wrong because its maxim allegedly cannot be universalized. Other acts of making false promises would have to be evaluated independently. Perhaps it will turn out that every act of making a false promise has a maxim that cannot be universalized. If so, CI_1 would imply that they are all wrong. So far, however, we have been given no reason to suppose that this is the case.

Other critics would insist that Kant hasn't even succeeded in showing that a_2 is morally wrong. They would claim that the first premise of the argument is false. Surely it could be a law of nature that everyone will make a false promise when in need of money, they would say. If people borrowed money on false promises rarely enough, and kept their word on other promises, then no contradiction would arise. There would then be no reason to suppose that "no one would believe he was being promised anything, but would laugh at utterances of this kind as empty shams."¹⁹

Let us turn, then, to the third example. Kant now illustrates the application of the categorical imperative to a case of imperfect duty to oneself. The action in question is the "neglect of natural talents." Kant apparently holds that it is wrong for a person to let all of his natural talents go to waste. Of course, if a person has several natural talents, he is not required to develop all of them. Perhaps Kant considers this to be an imperfect duty partly because a person has the freedom to select which talents he will develop and which he will allow to rust.

Kant imagines the case of someone who is comfortable as he is and who, out of laziness, contemplates performing the act, a_3 , of letting all his talents rust. His maxim in doing this would be:

$M(a_3)$: When I am comfortable as I am, I shall let my talents rust.

When generalized, the maxim becomes:

GM(a₃): Whenever anyone is comfortable as he is, he will let his talents rust.

Kant admits that *GM(a₃)* could be a law of nature. Thus, his argument in this case differs from the arguments he produced in the first two cases. Kant proceeds to outline the reasoning by which the agent would come to see that it would be wrong to perform *a₃*:

He then sees that a system of nature could indeed always subsist under such a universal law, although (like the South Sea Islanders) every man should let his talents rust and should be bent on devoting his life solely to idleness, indulgence, procreation, and, in a word, to enjoyment. Only he cannot possibly *will* that this should become a universal law of nature or should be implanted in us as such a law by a natural instinct. For as a rational being he necessarily wills that all his powers should be developed, since they serve him, and are given him, for all sorts of possible ends.²⁰

Once again, Kant's argument seems to be based on a rather dubious appeal to natural purposes. Allegedly, nature implanted our talents in us for all sorts of purposes. Hence, we necessarily will to develop them. If we also will to let them rust, we are willing both to develop them (as we must) and to refrain from developing them. Anyone who wills both of these things obviously wills inconsistently. Hence, the agent cannot consistently will that his talents rust. This, together with the categorical imperative, implies that it would be wrong to perform the act, *a₃*, of letting one's talents rust.

The argument can be put as follows:

Rusting-Talents Example

- (1) Everyone necessarily wills that all his talents be developed.
- (2) If everyone necessarily wills that all his talents be developed, then the agent of *a₃* cannot consistently will that *GM(a₃)* be a law of nature.
- (3) *a₃* is morally right if and only if the agent of *a₃* can consistently will that *GM(a₃)* be a law of nature.
- (4) Therefore *a₃* is not morally right.

This argument seems even less persuasive than the others. In the quoted passage Kant himself presents a counterexample to the first premise. The South Sea Islanders, according to Kant, do not will to develop their talents. This fact, if it is one, is surely inconsistent with the claim that we all necessarily will that all our talents be developed. Even if Kant is wrong about the South Sea Islanders, his first premise is still extremely implausible. Couldn't there be a rational person who, out of

idleness, simply does not will to develop his talents? If there could not be such a person, then what is the point of trying to show that we are under some specifically moral obligation to develop all our talents?

Once again, however, some philosophers may feel that Kant would have been worse off if his example had succeeded. These philosophers would hold that we in fact have no moral obligation to develop our talents. If Kant's theory had entailed that we have such an obligation, they would insist, then that would have shown that Kant's theory is defective.

In Kant's fourth illustration the categorical imperative is applied to an imperfect duty to others—the duty to help others who are in distress. Kant describes a man who is flourishing and who contemplates performing the act, *a₄*, of giving nothing to charity. His maxim is not stated by Kant in this passage, but it can probably be formulated as follows:

M(a₄): When I'm flourishing and others are in distress, I shall give nothing to charity.

When generalized, this maxim becomes:

GM(a₄): Whenever anyone is flourishing and others are in distress, he will give nothing to charity.

As in the other example of imperfect duty, Kant acknowledges that *GM(a₄)* could be a law of nature. Yet he claims once again that the agent cannot consistently will that it be a law of nature. He explains this by arguing as follows:

For a will which decided in this way would be in conflict with itself, since many a situation might arise in which the man needed love and sympathy from others, and in which, by such a law of nature sprung from his own will, he would rob himself of all hope of the help he wants for himself.²¹

Kant's point here seems to be this: The day may come when the agent is no longer flourishing. He may need charity from others. If that day does come, then he will find that he wills that others give him such aid. However, in willing that *GM(a₄)* be a law of nature, he has already willed that no one should give charitable aid to anyone. Hence, on that dark day, his will will contradict itself. Thus, he cannot consistently will that *GM(a₄)* be a law of nature. This being so, the categorical imperative entails that *a₄* is not right.

If this is Kant's reasoning, then his reasoning is defective. For we cannot infer from the fact that the person *may* someday want aid from others, that he in fact already is willing inconsistently when he wills today that no one should give aid to anyone. The main reason for this is that that dark day may not come, in which case no conflict will arise. Furthermore, as is pretty obvious upon reflection, even if that dark day does

arrive, the agent may steadfastly stick to his general policy. He may say, "I didn't help others when they were in need, and now that I'm in need I don't want any help from them." In this way, he would avoid having inconsistent policies. Unless this attitude is irrational, which it does not seem to be, Kant's fourth example is unsuccessful.

More Examples

It should be clear, then, that Kant has not provided us with a clear, persuasive example of the application of the categorical imperative. In light of this, some may feel that the categorical imperative is a worthless doctrine. Such a harsh judgment would probably be premature. For in the first place, Kant surely would have been worse off if he had succeeded in showing that suicide, or letting your talents rust, are invariably wrong. The normative status of these acts is hardly as obvious as Kant suggests. In the second place, the failure of Kant's illustrations may be due in part to his choice of some rather strange maxims, and to the fact that he presupposed some questionable views about the purposes of nature. Let us attempt to develop a more plausible illustration of the application of the categorical imperative.

In attempting to develop such an example, we should turn to the sort of case in which the categorical imperative stands the greatest chance of working correctly. This would be a case in which an agent proposes to take unfair advantage of his neighbors. It would be a case in which others, out of regard for the common good, have generously refrained from performing a certain kind of act, even though many of them might like to do such an act. Our agent, however, finds that he can get away with the act. The crucial feature of this case is that the agent cannot consistently will that the others act in the way he proposes to act. For if they all were to try to act in this way, that would destroy his opportunity for so acting.

Here is a good example of this sort of case. Primarily out of laziness, Miss Perkins, a college student, buys a term paper for her ethics course and submits it as her own work. Miss Perkins deals with a skillful term paper manufacturer, so she is assured of getting a very high grade. There is no chance that she will be found out. Most of us would say that regardless of its utility, Miss Perkins's act is morally wrong. She should not deceive her instructor and take advantage of her fellow students in this way. What does the categorical imperative say?

Let us call Miss Perkins's act of submitting the phoney term paper "a₅," and let us suppose that her maxim in performing a₅ is:

M(a₅): When I need a term paper for a course and don't feel like writing one, I shall buy a term paper and submit it as my own work.

The generalized form of her maxim is:

GM(a₅): Whenever anyone needs a term paper for a course and doesn't feel like writing one, she will buy one and submit it as her own work.

According to Kant's doctrine, a₅ is morally right only if Miss Perkins can consistently will that *GM*(a₅) be a law of nature. So to see if a₅ is right, we must determine whether Miss Perkins can consistently will that everyone needing a term paper but not feeling like writing one should submit a store-bought one.

It is reasonable to suppose that Miss Perkins cannot will that *GM*(a₅) be a law of nature. For consider what would happen if *GM*(a₅) were a law of nature, and everyone needing a term paper but not feeling like writing one were therefore to submit a store-bought one. Clearly, college instructors would soon realize that they were reading work not produced by their students. The instructors would have to deal with the problem—perhaps by resorting to a system under which each student would be required to take a final oral exam instead of submitting a term paper. If some such alteration in the course requirements were instituted, Miss Perkins would lose her opportunity to get a good grade by cheating. Thus, she surely does not will that any such change in the system should occur. She prefers to have the system remain as it is. Since it is clear that some such change would occur if *GM*(a₅) were a law of nature, Miss Perkins cannot consistently will that *GM*(a₅) be a law of nature. Thus, according to *CI*₁, her act is not right.

The essentials of this example are simple. Miss Perkins wills that the system remain as it is—thus providing her with the opportunity to take advantage of her instructor and her fellow students. She recognizes that if everyone were to submit a store-bought term paper, the system would be changed. Hence, she cannot consistently will that everyone should submit a store-bought term paper. In other words, she cannot consistently will that *GM*(a₅) be a law of nature. *CI*₁, together with this fact, entails that a₅ is morally wrong.

One of the most troubling aspects of this example is that it is pretty easy to see how the categorical imperative can be short-circuited. That is, it is pretty easy to see how Miss Perkins can make Kant's doctrine yield the result that her act is morally right. She needs only to change her maxim in a fairly trivial way:

$M(a_6)$: When I need a term paper for a course, and I don't feel like writing one, and no change in the system will occur if I submit a store-bought one, then I shall buy a term paper and submit it as my own work.

$M(a_6)$ differs from $M(a_5)$ in only one respect. $M(a_6)$ contains the extra phrase "and no change in the system will occur if I submit a store-bought one." But this little addition makes a big difference to the argument. We found that Miss Perkins could not consistently will that $GM(a_5)$ be a law of nature. For if she willed that $GM(a_5)$ be a law of nature, she would, indirectly, will that the system be changed. But she already willed that the system remain as it is. However, no such argument applies to $GM(a_6)$. For it appears that if $GM(a_6)$ were a law of nature, the system would not be changed. Apparently, then, Miss Perkins can consistently will that $GM(a_6)$ be a law of nature. Hence, according to CI_1 , her act of submitting a store-bought term paper, if performed under $M(a_6)$ rather than under $M(a_5)$, would be morally acceptable. This seems wrong.

The categorical imperative, interpreted as CI_1 , yields incorrect results in another sort of case too. Consider a man who has a large amount of money in a savings account. He decides that he will wait until the Stock Market Index reaches 1000 and then take all of his money out of the bank. This act seems quite acceptable from the moral point of view. However, it seems that CI_1 yields the odd result that the act is morally wrong. Let us consider why this is so.

We can call the man's act of removing his money from the bank "a₇." The maxim of a₇ is:

$M(a_7)$: When the Stock Market Index reaches 1000, I shall withdraw all my money from the bank.

The generalized form of $M(a_7)$ is:

$GM(a_7)$: Whenever the Stock Market Index reaches 1000, everyone shall withdraw all of their money from the bank.

It should be clear that the man cannot consistently will that $GM(a_7)$ be a law of nature. For banks have loaned out most of the money deposited in them. If everyone came to withdraw their savings from their bank, banks would soon run out of money. Not everyone can withdraw simultaneously. Hence, $GM(a_7)$ cannot be a law of nature. Thus, the agent cannot consistently will that it be so. CI_1 entails, together with this fact, that it would not be right for the man to withdraw his own money under this maxim. Surely, there is something wrong with a moral theory that has this result.

This same problem arises in any number of cases. Whenever, for some irrelevant reason, an otherwise innocent maxim cannot be universalized, CI_1 yields the result that the act is wrong. So if a person acts on the maxim, for example, of not becoming a doctor, he acts wrongly. For he surely could not will that *everyone* should refrain from becoming a doctor. As a rational being, he recognizes that there must be some doctors. Similarly, if a person acts on the maxim of always using adequate contraceptive devices when engaging in sexual intercourse, she acts wrongly, according to this interpretation of CI_1 . For if everyone were to do what she does, there would soon be no human race at all. This, Kant would think, is something no rational agent can consistently will.

These absurd results show that there is a very deep problem with CI_1 . The problem, in general, is that there are many different reasons why a maxim may fail to be universalizable. Some of these reasons have nothing whatever to do with morality. Yet, as far as can be discerned from the text of the *Groundwork*, Kant nowhere attempts to distinguish between innocent-but-nonuniversalizable maxims, on the one hand, and evil-and-nonuniversalizable ones, on the other. Without such a distinction, CI_1 yields obviously incorrect results in innumerable cases.

So we can conclude that there are very serious problems with CI_1 . Perhaps CI_1 is not an adequate interpretation of Kant's categorical imperative. Perhaps a more adequate version of that doctrine would not have these unsatisfactory results. However, if CI_1 is not Kant's theory, then it is very hard to see what Kant's theory might be.

NOTES

1. Immanuel Kant (1724–1804) is one of the greatest Continental philosophers. He produced quite a few philosophical works of major importance. The *Critique of Pure Reason* (1781) is perhaps his most famous work.

2. Kant's *Grundlegung zur Metaphysik der Sitten* (1785) has been translated into English many times. All references here are to Immanuel Kant, *Groundwork of the Metaphysic of Morals*, translated and analysed by H. J. Paton (New York: Harper & Row, 1964).

3. Kant, *Groundwork*, p. 70.

4. *Ibid.*, p. 88.

5. *Ibid.*, p. 89.

6. *Ibid.*, p. 91.

7. *Ibid.*, p. 69n.

8. In some unusual cases, it may accidentally happen that the situation to which the maxim applies can occur only once, as, for example, in the case of successful suicide. Nevertheless, the maxim is general in form.

9. Kant, *Groundwork*, p. 65.