Ethics // Spring 2019

Handout 4

Kant: Categorical imperative (FLN)

Categorical imperative. In following the categorical imperative we are supposed to abstract from purposes. Is the imperative simply of the form 'Act!'? This would be silly: though it does not specify a purpose, it should still say how we ought to act. Indeed, Kant explicates the 'no-purpose' requirement as leaving the agent with no leeway in his choice. That is, the HI really said two things: 420

> You ought to perform X, if you will to achieve Y and performing X is a means for achieving Y, but also: you do *not* have to perform X if either you do *not* will to achieve Y or performing X is *not* a means for achieving Y. (4-1)

If there is no external purpose to suggest to the agent how to act, then perhaps there is an internal feature of the imperative guiding the action. The internal feature left after purposes were removed is simply 'universality'. This means, Kant believes, that we have the following formula:

> *FUL* 'Act only on that maxim through which you can at the same time [rationally] will it should become a universal law.'

Having the formula FUL fills the gap observed earlier in the maxim of the moral shopkeeper.

Kant then proceeds to give three variants of this formula. How they are related we will briefly address later. The first variation, in any case, is:

> *FLN* 'Act *as if* the maxim of your action were to become through your will a universal law of nature.'

The task of CI is to provide a test for the maxims we adopt in our behaviour. CI specifies a necessary condition that a morally actionable maxim should satisfy. That is: a maxim is moral, and acting on it has moral worth, only if the maxim satisfies CI.

Kant works with four examples. Let us focus on the first two.

Suicide Suppose that, facing some serious problems (bankruptcy, unrequited love etc.), I decide to commit suicide. The reason, I notice, is self-love: I wish to avoid pain out of concern for myself (and not, say, out of concern for the overall amount of pain in the world). Then I in effect follow the maxim:

> Out of self-love I ought to shorten my life for the sake of avoiding future pain. (4-2)

If I test this maxim with FLN, then, Kant says, I see 'at once' that such a maxim could not become a universal law of nature. The reason is that self-love, by definition, is charged with preserving your life and at the same time is now also charged with terminating it. Contradiction.

Well, is it really? As a maxim of skill, it is not especially admirable, but perhaps there is nothing immoral about it either. Self-love may be charged with extending life in *pleasant* circumstances, rather than in *all* circumstances. I.e. self-love may have the task of increasing pleasures of the agent, rather than the task of preserving his life. It would be charged with preserving life only indirectly, since you can have pleasure only when you are alive. This is indeed a far more natural idea of self-love. Those people who committed suicide when captured by the enemy—they did it, we think, out of concern for themselves, and before that they went on with their lives out of the very same concern.

For Kant's reasoning to be convincing, he ought to show that my conduct must always be in accord with the rule: 'No one should ever commit suicide.' It is not at all clear that we can prove it by appeal to FLN (let alone FUL).

Remark 1. Another concern with the suicide example is that the teleological argument it is based on is rather feeble.

Question 2. Why does not Kant try to show the prohibition on suicide by the appeal to FUL?

Keeping promises Here the issue is about giving promises with the intention not to keep them. The maxim, yet again, has the trappings of a hypothetical form (since it refers to the goal of obtaining money), but can perhaps be restated in categorical form. In any case, if I were to universalise the

maxim, then I will see 'straight away' that in the world where everyone follows it, the practice of promises will disappear. Well, where is the contradiction? The reading I favour is that the maxim, now converted into a universal law, refers to giving promises, while at the same time its validity implies (in practice) the disappearance of any promise. I.e. the observance of this promise-creating law would lead to the disappearance of promises.

Puzzles. It has to be noted that many trivial maxims:

I will always wear blue hats (4-3)

and many ostensibly evil maxims:

I will insult Chinese people if I have an opportunity to do so (4-4)

can both pass the tests provided by FUL and FLN. There is no apparent contradiction involved in generalising them. The case of trivial maxims can be explained away by saying that FUL and FLN are meant to provide necessary conditions. That is, if a maxim is morally significant, then it must pass the generalisation tests. However, if it passes generalisation tests, this is no guarantee that it is morally significant. Another response is that for evaluating the moral value of maxims we must consider also their opposites. The maxim 'I will never wear blue hats' also passes the generalisation tests. Both of them are permissible, and because of that, both of them are also morally neutral.

It is more difficult to see what is wrong with ostensibly evil, yet universalisable, maxims. One way out here is to search for a more basic maxim underlying my more specific maxim and then try to fail it with FUL and FLN. In this instance it will be:

I will insult people. (4-5)

Finally, there are maxims that are ostensibly innocent, yet not universalisable. For instance, the maxim:

I will buy gold and never sell it. (4-6)

This is expected: there may be logical, or economical, or psychological, or physical inconsistencies that are not *moral* inconsistencies. So we have maxims that do not violate any *moral* law, but still cannot be coherently willed as universal laws.

The apriority trap. It is not clear how seriously, from Kant's point of view, we should take those puzzles, or indeed, *any* puzzles. If Kant intends his moral theory to be *a priori*, then the theory should be immune from any kind of empirical justification or refutation. So, e.g., even if the maxim (4-3) does not seem to us immoral, the Kantian theory should not be compelled to take account of this fact. Our intuition about (4-3) is just that, an intuition. Kant has already officially refused to invoke intuitions in support of moral theory. Why, therefore, should he worry that these intuitions have any say in morality? Moral theory, as Kant understands it (and as I understand him), is not an empirical theory tasked with delivering accurate predictions or explaining our antecedent beliefs. If it were, then of course the puzzles above would be urgent to resolve.

An analogy here is with mathematics. Suppose it is a mathematical law that $2 + 2 = 4$. And suppose that every time you put two objects and two objects together you get five when counting them (think of rabbits in a cage). We shouldn't think that this fact has *any* relevance for mathematics. Instead we say either that your counting is impaired, or that there is a weird physical irregularity in place. Our mathematical law is, in other words, immune to any empirical data.

But of course, this is a sign of trouble. One might accuse Kant of inventing an arbitrary doctrine of morality (call it 'K-morality') that does not necessarily, or not all, match what *we* understand by 'morality'. The verdicts of one do not match the verdicts of the other, and they do not have to do so.

Hence a dilemma: *Either* we have a pre-theoretic conception of morality that we attempt to set on a sound basis. This conception is equipped with all sorts of moral judgements that cannot *entirely* be abandoned. In this case, our moral theory cannot be wholly *a priori*, and generally speaking, it will be sensitive to the kind of puzzles we had above. *Or* else we come to examine a moral theory with no moral judgements at all (since we reject their validity anyway). Then we can have no puzzles grounded in our pre-theoretic intuitions. And the trouble is, we then simply invent K-morality that has no relevance to the familiar moral problems we struggle with in our non-philosophical lives.