Obligatory actions, obligatory maxims

Abstract: In this paper, I confront Parfit's Mixed Maxims Objection. I argue that recent attempts to respond to this objection fail, and I argue that their failure is compounded by the failure of recent attempts to show how the Formula of Universal Law can be used to demarcate the category of obligatory maxims. I then set out my own response to the objection, drawing on remarks from Kant's *Metaphysics of Morals* for inspiration and developing a novel account of how the Formula of Universal Law can be employed to determine the deontic status of action tokens, action types, and maxims.

Keywords: Kant's ethics, Kantian ethics, Parfit, Mixed Maxims Objection, On What Matters, Formula of Universal Law, positive duty, general duty, obligation, action description problem

In *On What Matters*, Derek Parfit questions whether Kant's Formula of Universal Law (FUL) can provide agents with action guidance. The question arises because the FUL is in terms of maxims and there is a many-many relation between maxims and actions: even if the FUL is able satisfactorily to determine the deontic status of any given maxim, that has no obvious bearing on the deontic status of any given action. According to Parfit's Mixed Maxims Objection, this has 'unacceptable' results (Parfit 2011: 290).

In this paper I argue that recent attempts to deal with the Mixed Maxims Objection fail, in part because they fail to appreciate its scope. Parfit illustrates the objection with intuitively permissible actions that the FUL seems to render impermissible, and therefore Kantians have concentrated on the relationship between permissible maxims and permissible actions. But in so doing, they overlook other deontic categories, like the obligatory. This problem is then further compounded by the failure of recent and independent attempts to derive obligatory maxims from the FUL. I want to show that there is a solution to the problem Parfit raises, but I want to show that the solution is different from what recent philosophers have taken it to be.

The paper is divided into four sections. In the first, I introduce Kant's FUL and Parfit's Mixed Maxims Objection. In the second, I explain and then criticize recent responses to Parfit's Mixed Maxims Objection. In the third, I criticize recent attempts to show how the FUL can be used to determine the deontic status of maxims. In the fourth, I build on Kant's *Metaphysics of Morals* to sketch my own positive solution to these issues.

Section 1. Parfit's Mixed Maxims Objection

Kant's FUL is articulated first in part I of the *Groundwork for a Metaphysics of Morals*. It runs as follows: 'I should never proceed otherwise than thusly, that I also could will my maxim should become a universal law' (G, 4: 402.07-09).[1] From this it may be seen that the FUL is a moral test of maxims. If a maxim is universalizable, then it is permissible. If a maxim is not universalizable, then it is impermissible.

Kant gives various examples of impermissible maxims. One of the most famous is a suicide maxim: 'I make it into a principle for myself, when, with a longer term, life threatens more evil than it promises pleasure, to shorten it for myself from self love' (G, 4: 422.04-07). Another of the most famous is a false promise maxim: 'When I believe myself to be in financial distress, then I will borrow money and promise to repay it, even though I at the same time know that it will never happen (G, 4: 422.22-24).

In discussing both of these maxims, Kant follows a similar procedure. He gives a vignette in which an agent is unsure about how to proceed, and he then explains why these maxims are immoral. One

---

[1] Special typescripts removed. All translations in this paper are my own. Citations are in accordance with the standard Academy Pagination, which can be found in the margins of many translations. I use the following abbreviations: G-*Groundwork for a Metaphysics of Morals*; and MM—*Metaphysics of Morals*.

thing that Kant does not do, unfortunately, is give a detailed account of what exactly a maxim is. This has given rise to debate.[2] But the following three claims about maxims seem to be accepted on all sides:

(1) A maxim is an agent's principle of volition.
(2) An agent can act on the same maxim on more than one occasion.
(3) A maxim can incorporate more than (merely) an act description.

(1) will become more important in section three of this paper. For now I want to focus on (2) and (3).

The idea behind (2) can be illustrated by example. Suppose that an agent has adopted a maxim of beneficence. Then the agent might act on this maxim on one occasion when helping out at a soup kitchen, and she might act on the same maxim on another occasion when helping a friend to move. Along the same lines, an agent who adopts Kant's false promise maxim might act on it on one occasion when trying to get money to pay rent, and she might act on the same maxim on another occasion when trying to get money to pay for groceries. From this it may be seen that a single maxim can be instantiated in many act tokens.

Some might object at this point on the grounds that an agent cannot commit suicide more than once and so Kant's suicide maxim does not satisfy (2). This objection poses a dilemma: either (2) is incorrect, or the suicide maxim must be explained away as nonstandard.

However, I do not think that this objection works. The problem with this objection is that it rests on a false presupposition, that an agent who acts on a maxim will be successful. But just as an agent might be caught in a lie and therefore fail in her efforts to get ready money with the false promise maxim, and just as an agent accidentally might burn a soup kitchen down while trying to help prepare food and therefore fail in her efforts to help with a maxim of beneficence, so an agent might botch a suicide attempt. And as much as we might not want an agent who botches a suicide attempt to try again, she very well might do so, acting on the same suicide maxim on more than one occasion.

The idea behind (3) also can be illustrated by example. A standard maxim of beneficence used in the secondary literature is 'to help others'. But most Kantians accept that this maxim also could contain the general conditions under which the agent would act, as in 'to help others if they are in great need'.[3] Similarly, the suicide maxim contains the agent's motive ('from self love'), and the false promise maxim contains a description of the agent's circumstances ('in financial distress'). From this it may be seen that a maxim can contain more than merely an action type and, therefore, that there is more than one maxim for any given action type.

Now an act token can instantiate multiple action types.[4] In conjunction with (2) and (3), this then seems to entail that there are many-many relations between maxims and action types and between maxims and action tokens, and these many-many relations, in turn, give rise to Parfit's Mixed Maxims Objection. Let me explain.

According to Parfit, 'many maxims are *morally mixed* in the sense that, if we always acted on these maxims, some of our acts would be wrong, but other acts would be permissible or even morally required' (Parfit 2011: 293). For example, Parfit considers an egoist's maxim, 'Do whatever would be best for me'. This maxim is impermissible according to the FUL. But as Parfit points out, not all actions an egoist would perform on this maxim are intuitively impermissible. If the egoist were to engage in acts of larceny on this maxim, his actions would be impermissible. But when 'for self-interested reasons, he pays his debts, keeps his promises, and saves a drowning child, because he hopes to get some reward', such actions are also condemned by the FUL because of the maxim underlying them (Parfit 2011: 290). This, Parfit claims, is counterintuitive.

Parfit considers the possibility that it is not the agent's action that is wrong but rather the agent's doing of the action that is wrong. But Parfit rejects this. Parfit points out that when the egoist puts on some warmer clothing, '[t]here is no sense in which...his acting in this way is wrong' (Parfit 2011: 291).

_____

[2] For a summary, see (Gressis 2010a) or (2010b).

[3] Some argue that a well-formed maxim always will have the form 'to A if C', where A is an action description and C is the set of conditions under which the agent would perform A.

[4] This point, raised to prominence in Elizabeth Anscombe (1957), generates the problem of relevant act descriptions, arguably a forerunner of Parfit's Mixed Maxims Objection.

Thus Parfit concludes that the deontic status of an agent's maxim is often irrelevant to the deontic status of an agent's action:

> When we ask whether some act is wrong, or contrary to duty, Kant's formula often makes the answer depend on morally *irrelevant* facts. When my Egoist risks his life to save some drowning child, it is irrelevant that he is acting on the policy of doing whatever would be best for himself. When Kant told someone the correct time, it was irrelevant that he was acting on the [non-universalizable (according to Parfit)] policy 'Never lie'. These facts at most give us reasons to believe that in some *other* cases this Egoist would, and Kant might act, wrongly. (Parfit 2011: 295)

And if this is correct, then the FUL fails to categorize act tokens correctly; it provides no information with regard to act types; and it fails to provide action guidance and to help third-party observers to determine the deontic status of an action.

Section 2. Recent responses to Parfit

Many Kantians do not accept these negative conclusions about the FUL.

Thomas Pogge concedes that '[t]he Categorical Imperative is a criterion for the permissibility of maxims, and Kant intends this criterion to play a subsidiary role in the assessment of character ('good will')—not in the assessment of act tokens' (Pogge 2004: 54). Here it might seem like Pogge is saying that Parfit is right but that Parfit's criticism is misguided because the FUL is not intended to provide action guidance.[5] But Pogge then argues:

> To be sure, Kant held beliefs not only about his questions: 'when is a maxim morally wrong?' and 'When does a person have a good will?', but also about Parfit's question: 'When is an act token morally wrong?'. But Kant does not provide a clear path from the first to the last question. The path cannot be this: An act token is morally wrong just in case it is performed on an impermissible maxim...The path must be something like this: An act token is wrong (contrary to duty) just in case any maxim on which it might be performed is impermissible. (Pogge 2004: 55)

That is, according to Pogge, an act token is morally wrong if but only if all maxims on which it might be performed are impermissible.[6]

Applying Pogge's criterion to Parfit's egoist, it may be concluded that although the egoist's maxim is wrong and the egoist does not have a good will, many of the egoist's actions are permissible because they *could* be performed on permissible maxims (even though in fact they are not). Thus Pogge solves Parfit's Mixed Maxims Objection by severing the link between the deontic status of an agent's maxim and her action while nonetheless preserving the relevance of maxims *in general* to the deontic status of an action.

The same account is advanced independently by Sven Nyholm:

> The actions of Parfit's egoist in his examples (saving a drowning child, keeping a promise) are, in Kant's sense, actions that can coexist with autonomy, since they are actions that somebody governed by maxims fit to be universal laws could perform in the circumstances on the basis of operating on such principles, thereby exercising what Kant calls autonomy. (Nyholm 2015: 287)

---

[5] Jens Timmermann, writing before Parfit's Mixed Maxim objection, seems to subscribe to a position like this: 'the categorical imperative cannot serve as a standard to determine the moral rightness of specific acts' (Timmermann 2000: 50). As will emerge in section 4 below, I disagree.

[6] Although Pogge does not mention it, I think he must be assuming a veridicality clause. That is, I do not think Pogge would take it as a problem if an agent could perform an impermissible action on a permissible maxim only because the agent has mistaken beliefs about what ze is doing.

According to Nyholm, an action is permissible if but only if there is some permissible maxim on the basis of which it can be performed. Because X (an action or a maxim) is permissible if but only it is not impermissible, Pogge's and Nyholm's accounts are formally equivalent.

I would like to raise two problems for Pogge and Nyholm.

First, their account makes it in principle impossible to determine whether a given action is impermissible. Let me explain.

Pogge and Nyholm (along with Parfit) presuppose that act tokens and maxims can be separated.[7] But if that is so, then there are potentially infinitely many maxims on which any given act token might be performed. It follows immediately that, in order to determine whether a given action is impermissible, Pogge and Nyholm require us to test infinitely many maxims. But no human agent ever will be able to test infinitely many maxims, so no human agent ever will be able to determine whether a given action is impermissible.

To be clear: it might be possible to determine whether a given action is permissible. To determine whether an act token is permissible might require testing only one maxim.[8]

But the permissible is the only category that can be applied to actions on this account, and it can be applied only if an agent is lucky enough to test a permissible maxim on which the action in question can be performed.[9] If an agent has tested only impermissible maxims on which an action can be performed, she will not be in a position to determine whether the action is permissible or impermissible. This, I think, is a problem for an account that attempts to explain how the FUL can be used to determine the deontic status of actions.

Second, Pogge and Nyholm overlook deontic categories like the obligatory. The only categories Pogge and Nyholm mention are the permissible and the impermissible. This is understandable given that Parfit's egoist example focuses on these two categories. But it is, I think, a major lacuna.

Here are four ways in which Pogge and Nyholm might try to extend their account to deal with this second problem:

1. Define the obligatory by exclusion: an action is obligatory if but only if it is neither merely permissible nor impermissible.
2. Define the obligatory in terms of the impermissible: X is impermissible if but only if not-X is obligatory.
3. Define the obligatory as Pogge defines the impermissible: an action is obligatory if but only if all maxims on which it can be performed are obligatory.

_____

[7] This follows immediately from the fact that all three of them accept that (a) the egoist's act tokens are entirely permissible even though (b) the maxims on which the egoist is acting are impermissible. Pogge and Nyholm presuppose, further, that an act token can be performed on a maxim other than the one on which it actually is performed (they maintain that the egoist's act tokens are permissible because the egoist's action tokens *could* have been performed on permissible maxims even though in fact they are not).

An alternative (and more complicated) account of how to assess an act token separately from the maxim on which it is performed could be based on the conjunction of the following two biconditionals: (i) an act token is permissible if but only if the act type in the maxim on which it is performed is permissible; and (ii) an act type is permissible if but only if it is possible to incorporate it into a permissible maxim. From (i) and (ii) it follows that an act token can be permissible even if it is performed on an impermissible maxim, for there might be a permissible maxim that incorporates that same act type. For example, because the act types in the egoist's maxims (putting on warm clothing, saving a child, etc.) could be put into permissible maxims, (i) and (ii) entail that the egoist's act tokens are permissible even though the egoist's maxims are not.

I want to say two things about this alternative account.

First, I want to emphasize that this is not the account advanced by Pogge and Nyholm.

Second, this account would have absurd consequences. To see why, note that most impermissible act types (like lying) can be incorporated into permissible maxims (something Kant himself seems to suggest in the casuistical questions in the *Metaphysics of Morals*). This, conjoined with (i) and (ii), entails that act tokens of these otherwise impermissible actions are *always* permissible. So I think that this alternative account can be dismissed fairly straightforwardly. As will emerge in section 4 of this paper, I reject both (i) and (ii).

[8] For this reason Nyholm's account might be preferred over Pogge's on pragmatic grounds. Although the two accounts are formally equivalent, Nyholm's might be thought to make more perspicuous that to determine whether an act token is *permissible* only a finite number of maxims must be tested.

[9] It is worth pointing out that there is no fixed enumeration of maxims. So we do not have even a semi-decision procedure for determining the permissibility of action tokens. But overlooking these formal issues, the fact remains that given Pogge's and Nyholm's account, it is in principle impossible for an agent to determine whether a given action is impermissible.

4. Define the obligatory as Nyholm defines the permissible: an action is obligatory if but only if there is some obligatory maxim on the basis of which it can be performed.

I am now going to show that all four of these are flawed.

There are at least three problems with 1. First, Nyholm and Pogge do not offer an account of the *merely* permissible. They offer an account of the permissible *sans phrase*, and they accept (what is an axiom of deontic logic) that the obligatory is permissible. So attempting 1 requires, in turn, retooling Nyholm and Pogge.

Second, even if Nyholm's and Pogge's accounts could be retooled in terms of mere permissibility, defining the obligatory by exclusion would inherit the first problem above regarding the undecidability of impermissibility.

It might be thought that a more complicated definition by exclusion could be used to avoid this second problem: an action is obligatory if but only if it is permissible but not merely permissible. But this plays into the third problem I want to raise: defining the obligatory in this way rules out other deontic categories, like the supererogatory and the suberogatory. These categories might turn out to be conceptually confused and, thus, empty. But it would be better not to adopt an account of obligation that rules them out from the start.

There are at least two problems with 2. First, like 1, 2 inherits the problem the original account has with the impermissible.

Second, even if 2 were supplemented with an account of when act tokens are impermissible (to patch the first problem), further supplementation would be required to determine when an act token is obligatory. This is for the simple reason that there are infinitely many action tokens that satisfy the description 'not-X' in the biconditional above, and it is both counterintuitive and contrary to widely accepted axioms of deontic logic to say that all of these actions are obligatory. From this it may be seen that 2 fails on its own terms.

The problem with 3 is like the problem with the impermissible. Because 3 requires testing an infinite number of maxims, there is in principle no way to determine whether a given action is obligatory.[10]

This last problem is obviated by 4. But if it is obligatory to adopt the maxim 'always to act in conformity with the moral law', then 4 renders all permissible actions obligatory. Thus, 4 risks eliminating the category of the merely permissible with regard to actions.

However, the problems that I have raised against 3 and 4 might seem, to some, less pressing than the problems I have raised against 1 and 2. They might think that 3 or 4 can be fixed. Others simply might not accept the problems that I have raised against 3 or 4. For that reason, in the next section of this paper I am going to address what I take to be the deepest problem with 3 and 4: the FUL cannot give an account of obligatory maxims.

Section 3. Deriving positive duties from the FUL

The FUL seems to divide maxims into two kinds: those that are universalizable and those that are not. The former are permissible whereas the latter are impermissible. There is then a question about how to determine which maxims are obligatory.

To deal with this, Kantians apply the FUL to pairs of maxims. Given a maxim 'to B', there is then a corresponding maxim of omission/rejection/negation, the maxim 'not to B'. If one of these is universalizable but the other is not, then the universalizable one is claimed to be obligatory.

I am going to examine two ways in which this argument has been defended. The first appeals to logical necessity; the second appeals to psychological necessity. I maintain that both fail.

---

[10] If it was not clear from the above, this is not because one might be uncertain about the deontic status of a given maxim. It is because finite human agents only can test a finite number of maxims during their lifetimes. So even if they are certain of the deontic status of every maxim they test (and never forget any result), they never will be warranted in concluding that a given action is obligatory on 2.

Subsection 3.1 Logical necessity

The maxim 'to B' and the maxim 'not to B' are sometimes taken to be the maxim-analog of contradictory propositions. If two propositions are contradictories, then the truth of one entails the falsity of the other. Similarly, given two contradictory maxims, if an agent has adopted one, this (logically) entails that he has not adopted the other. Thus, if a maxim is permissible but its contradictory is not, then the maxim is obligatory.

This argument seems to be made by Paul Guyer, Jonathan Harrison, Barbara Herman, Onora O'Neill, and Philip Stratton-Lake:

> [J]ust as the falsity of one statement implies the truth of its contradictory, e.g. 'It is false that it is raining today' implies 'It is true that it is not raining today,' so it would seem that the impermissibility of adopting one maxim does imply the necessity of adopting its contradictory, thus the impermissibility of the maxim of never helping anyone does imply the necessity of the maxim of helping somebody sometime. (Guyer 2007: 143)

> Kant holds (a) that a maxim is not morally acceptable and must not be adopted...if it cannot be universalized; (b) that it may be adopted...if it can be universalized. He may also have thought (c) that a maxim must be adopted (ought to be adopted) if what I shall loosely call its 'contradictory' is not universifiable...(c) would follow from (a) together with the additional premise, which I see no reason to cavil about, that if it is our duty not to do A, then it is our duty to do non-A. (Harrison 1957: 22)

> ...if the CI procedure shows that it is impermissible to adopt the maxim, 'to never help anyone,' it follows that we must adopt its contradictory, 'to help some others sometimes.' (Herman 1993: 63)[11]

> An agent choosing a policy with respect to *x* must have one, but only one, of each of the pairs of contradictory maxims. (O'Neill 2013: 181)

> ...a principle is morally required if it is the contradictory of a maxim that cannot be conceived or willed as a universal law. (Stratton-Lake 2000: 88)[12]

The central premise in this argument is that all agents adopt either a maxim ('to B') or its contradictory ('not to B'). As the first two quotations (from Guyer and Harrison) reveal, this premise seems to be motivated by the idea that either an agent will B or he will not B, from which it is supposed to follow that either an agent has a maxim 'to B' or a maxim 'not to B'. I would like to say two things about this.

First, this argument runs afoul of the claim, noted in section 1 and accepted on all sides, that maxims are agents' volitional principles. Let me explain.

Kant thinks that what distinguishes rational from nonrational beings is that the former are able to act in accordance with the representation of a law:

> The will is thought as a capacity to determine oneself to action *according to the representation of certain laws*. And such a capacity can be encountered only in rational nature. (G, 4: 427.19-21)

That is, whereas both rational and nonrational beings will conform to laws of nature like the law of gravity, only rational beings will govern their actions in accordance with principles of volition. These

---

[11] Herman makes a similar argument in her (2007: 208).

[12] Stratton-Lake also applies this line of reasoning to actions: '...an action is morally required if its maxim is universalizable while the maxim of the contradictory act is not' (2000: 88).

principles are maxims. The point is that the inference from *either an agent will B or he will not B* to *either an agent has the maxim 'to B' or he has the maxim 'not to B'* involves an illicit move from external descriptions of an agent to internal principles of volition, one that will go through only if it is supplemented with a conception of maxims that contradicts the claim that maxims are agents' volitional principles.

Second, this account entails that all agents have infinitely many maxims, a *prima facie* implausible result given our finite capacities.

That is, this account entails that for any action B, all agents have the maxim 'to B' or 'not to B', and there are infinitely many actions that can be plugged into this. Not only is it implausible that *all* agents have adopted *some* such maxims (e.g., it is implausible that 18th century agents living in the forests of Brazil had adopted either the maxim 'to assassinate Hitler' or the maxim 'not to assassinate Hitler'), but more: it is implausible that *any* (human) agents have adopted *all* such maxims. The appeal to logical necessity, however, entails both of these results, for the appeal to logical necessity rests on the claim that *all* agents have adopted one or the other of *all* such maxim pairs. And given the foregoing, this, I suggest, is unsustainable.[13]

Subsection 3.2 Psychological necessity

An alternate route to obligatory maxims appeals to psychological necessity. Instead of conceiving of 'to B' and 'not to B' as the maxim-analog of contradictory propositions, they are conceived as the maxim-analog of contrary propositions. If 'to B' and 'not to B' are contraries, then no agent will adopt both (at the same time). But some might adopt neither.

It is then asserted that when an agent (like Hamlet) is engaged in deliberation about B-ing, she must adopt one of these two maxims (assuming she does not fall asleep or something before concluding her deliberations). Thus, if a maxim is permissible and its contrary is not, then the maxim is obligatory *for any agent engaged in deliberation* about the action in question.

Because some actions (e.g., helping others, keeping promises, etc.) are so pervasive to the human condition, all normal adults will have engaged in deliberation about them at some point in their lives. So for pragmatic reasons the deliberation-caveat may be dropped in some cases.

This seems to be the strategy appealed to by Fabian Freyenhagen and Jill Hernandez:

> A strict duty is one where universalizing a maxim contrary to it would generate a contradiction in conception, while a wide duty is one where universalizing a maxim contrary to it would avoid this problem, but instead generate a contradiction in willing. (Freyenhagen 2012: 56)

> Kantian ethicists typically classify actions in three ways: morally permissible acts (whose maxim and its contrary do not contradict the moral law), morally obligatory acts (whose contrary maxim conflicts with the moral law), and morally forbidden acts (whose maxim contradicts the moral law). (Hernandez 2010: 403)[14]

Guyer also sometimes adopts this strategy:

---

[13] To be clear: all agents should have maxims that would entail, e.g., that they would not help out at Auschwitz-Birkenau if it were still running. But that does not entail that they will have maxims about the concentrations camps. Relatedly, it is worth remembering that there is a difference between maxim-based omission of an action and omission of an action. For example, an agent can omit to rape deliberately, as when Svidrigailov deliberately refrains from forcing himself on Dunya. But agents also can not rape other agents without adopting a maxim not to do so. Most agents do not commit rape while they are asleep, and (in general) sleep-based omissions can but need not be maxim governed.

[14] Hernandez mixes talk of maxims and actions here, and the content of her article suggests that she does not realize there is a distinction. Although (in section 4 of this paper) I shall reject the separation of maxims from act tokens, I want to be clear that there nonetheless would remain (on my account) a many-many relation between maxims and act tokens, for a single maxim can be acted on more than once, and a given act token might manifest multiple of an agent's maxims.

> ...if I must reject the maxims of letting all my talents rust or never helping anyone else, then I must accept their logical contraries, namely, maxims of cultivating at least some of my talents and helping at least some other people some of the time. (Guyer 2006: 194)

Similarly, although O'Neill appeals to maxim contradictories in the quotation above (when talking about 'maxims of ends'), at other places she appeals to maxim contraries (when talking about 'maxims of action').[15] She argues that if the maxim 'to A if B' yields no contradiction when universalized, then the FUL must be applied to the contrary of this maxim, 'to omit A if B':

> If the test now yields a contradiction, then it is forbidden to omit A if B and consequently obligatory to do A if B. (O'Neill 2013: 163)

The same argument is also made by Christine Korsgaard, although she does not appeal to the concept of a maxim contrary explicitly:

> Kant does derive obligatory ends from the Formula of Universal Law, but he does it by a curiously roundabout procedure in which someone is imagined formulating a maxim of rejecting them and then finding it to be impermissible. This argument does not show that there would be a moral failing if the agent merely unthinkingly neglected rather than rejected these ends. The point about the pervasiveness of these ends in the moral life is a more complicated one, one that follows from their adoption by this route… (Korsgaard, 1996: 152-153)

I would like to say two things about this argument.

First, the central premise of this argument, that an agent deliberating about B-ing will adopt either the maxim 'to B' or its contrary, the maxim 'not to B', is false.

An agent engaged in deliberation about B-ing will face a multitude of maxims, and they need not be contraries to be deliberative options. In addition to the maxims 'to B' and 'not to B', an agent might adopt 'to B if C', 'to B if D', 'to B if E',..., 'not to B if C', 'not to B if D', 'not to B if E'.

To make this more concrete, an agent deliberating about whether to help someone else in need might adopt any of the following maxims:

1. To help others.
2. To help friends.
3. To help family.
4. To help members of my club.
5. Not to help anyone.
6. Not to help friends.
7. Not to help family.
8. Not to help members of my club.

A proponent of this argument strategy might contend that the argument can be retooled to meet this objection. Instead of testing only two maxims ('to B' and 'not to B'), test all of the deliberative

---

[15] It should be noted that O'Neill does not define what she means by a 'maxim contradictory'. But she does define what she means by a 'maxim contrary,' and her definition is different from mine. Here is O'Neill's definition: 'I shall call the pair of maxims which must be tested in any complete application of the contradiction in conception test *contraries*, because their UTC's are contraries, and not contradictories' (O'Neill 2013: 162n25).

The UTC of a maxim is its Universalized Typified Counterpart. Thus, the UTC of 'I will A if B' is 'everyone will A if B', where the 'will' of the first is that of practical reason whereas the 'will' of the second is that of a law of nature. (As O'Neill is careful to point out in the footnote from which I have taken her definition, the UTC is a proposition whereas a maxim is not, thus the former has a contrary whereas, strictly speaking, the latter does not. That is why she defines the concept of a maxim contrary so as to make it derivative of the more conventional concept of a propositional contrary.)

From this it may be seen that O'Neill's definition and mine are formally equivalent (in the sense that any two maxims that are contraries on her definition will be so on mine and *vice versa*).

options. If all but one of the deliberative options are impermissible, then the agent has a positive duty to adopt the only permissible maxim. This leads me to my second point.

My second point is that the list of helping maxims has two notable properties. One is that lists of this nature can be continued *ad infinitum*. Thus this argument strategy makes it impossible for us to use the obligatory in the same way that Pogge and Nyholm make it impossible for us to use the impermissible. The other is that it is *prima facie* implausible that only one of the maxims on this list will be permissible.

Alternatively, a proponent of this argument strategy might maintain that if an agent adopts a maxim 'to B if C' (for any C), then she *eo ipso* has adopted the maxim 'to B', and if an agent adopts a maxim 'not to B', then she *eo ipso* has adopted the maxim 'not to B if C' (for any C). Thus if the maxim 'to B' is permissible but the maxim 'not to B' is impermissible, the former is obligatory regardless of whether an agent might adopt the maxim 'to B if C' and some such maxim is permissible. However, there are two problems with this.

The first problem is that the argument rests on the unstated premise that if 'not to B' is impermissible, then all maxims of the form 'not to B if C' are impermissible. But it is *prima facie* implausible that all such maxims will be impermissible ('not to help criminals', 'not to help Nazis', and 'not to help skinheads' serve to illustrate this latter point when the action is helping).

The second problem is that the claim that an agent who adopts the maxim 'to B if C' *eo ipso* adopts the maxim 'to B' rests on ideas that (again) contradict the fact that maxims are agents' volitional principles. For example, this claim might rest on a maxim entailment principle according to which an agent counts as having adopted a maxim M if she has adopted some other maxim N and the proposition corresponding to N entails the proposition corresponding to M. But such a principle would face the same problems that face the idea of maxim contradictories: opacity considerations; infinite maxims with finite capacities; etc.[16] Moreover (and regardless of the point made in the previous sentence), an agent who adopts the maxim 'to help family' might balk at having the associated general maxim of helping ascribed to him: unspecified as it is, the maxim is too unrealistic to constitute a genuine deliberative option or volitional principle.

Section 4. An alternate route

In the previous two sections of this paper I argued for two related conclusions: (1) existing attempts from Kantian ethicists to respond to Parfit and elucidate the relationship between permissible maxims and permissible actions fail, and (2) existing attempts from Kantian ethicists to derive obligatory maxims from the FUL fail.

In this section I am going to sketch my own solutions to these problems. First I am going to argue for an alternative relationship between permissible maxims and permissible actions. Then I am going to use that relationship, along with some remarks Kant makes in the *Metaphysics of Morals*, to explain how to account for the intuitions motivating the arguments discussed in section 3. Doing so also will help to show how obligatory actions can be derived from the universalizability tests even though obligatory maxims cannot.

I maintain that an act token is permissible if but only if the maxim on which it is performed is permissible.[17] This avoids the problems associated with Nyholm and Pogge because it does not require testing all possible maxims on which an action might be performed. It requires testing only the actual maxim on which an action actually was, is, or will be performed, where the maxim on which an action is performed is the principle on the basis of which an agent actually acts.[18]

---

[16] Indeed, the fact that such an entailment principle requires distinguishing between two different kinds of maxim adoption (regular adoption and 'adoption by entailment') even in its formulation already suggests that something has gone awry.

[17] Complications can arise here if an act token instantiates multiple maxims and those maxims have conflicting deontic statuses. However, I cannot pursue such complications here.

[18] It might help to think of the Aristotelian notion of a practical syllogism; the maxim on which an action is performed is the 'major premise' that leads to the action as 'conclusion'.

The problem with this solution is that it seems to entail Parfit's 'unacceptable' result, that his egoist is acting impermissibly when putting on warm clothing on the basis of the impermissible maxim 'Do whatever would be best for me'. So let me explain how I avoid this problem.

Parfit does not consider that his egoist will have subordinate maxims. The egoist might motivate the adoption of those subordinate maxims on the basis of an egoistic principle. But it is those subordinate maxims that are manifested in the action of putting on warm clothing. And it is the maxim that immediately motivates the action that, I am suggesting, must be tested in order to determine the permissibility of the action.

For example, an egoist who puts on warm clothing on the basis of the maxim 'to keep warm' is acting permissibly. An egoist who puts on warm clothing on the basis of the maxim 'not to lend unworn clothing to those in need' is acting impermissibly.

The reason Parfit and his respondents do not consider this solution has to do with what Parfit calls the 'Rarity Objection' (Parfit 2011: 289, emphasis omitted). According to this objection, if enough detail is included in any maxim, that maxim will pass the FUL. On the basis of this objection it may be inferred that maxims should not include too much detail.

I am not convinced that the Rarity Objection succeeds: I do not see why including detail would entail that a maxim will be universalizable. But I do not want to enter into debate about how to interpret the FUL properly, and I do not need to in order to answer the Rarity Objection. Regardless of whether including enough detail will render any maxim universalizable, there is a gap between ruling out highly specific and detailed maxims and concluding that only a highly general principle like the egoist's counts as a maxim. So regardless of whether the Rarity Objection succeeds, it does not show that the egoist's more specific principles are not subject to moral assessment.[19]

Parfit *et al.* might object that I have sidestepped rather than addressed the Mixed Maxims Objection. They might contend that in addressing the particular example used to illustrate the problem, I have lost sight of the underlying issue: that (a) maxims can manifest in multiple act tokens and (b) the deontic status of the maxim remains constant (according to the FUL) even though the deontic status of the act tokens does not.

But any such objection is based on a twofold error. First, it confuses act tokens with act types and, second, it rests on the separability of act tokens and their associated maxims.

The claim that act tokens can be separated from their associated maxims itself rests, I suspect, on a confusion of act tokens with act types. Although a given act type can be performed on multiple different maxims, a given act token cannot be.

To be sure, a maxim can manifest in multiple act tokens. Thus there is a one-many relationship between maxims and act tokens. Similarly, many maxims can manifest in a single act token, whence it follows that there is, in some sense, a many-one relationship between maxims and act tokens.

But the self-same act token cannot be performed on varying maxims: if an agent saves a drowning child from considerations of humanity whereas her counterpart saves the same drowning child from considerations of cannibalism, they are performing different act tokens even though their act tokens might instantiate some of the same act types.

From this it follows that identifying the maxim on which an act token is performed is necessary but not sufficient to identify the act token. But identifying the deontic status of the associated maxim is, I contend, sufficient for identifying the deontic status of the act token, for any other identifying information (not included in the agent's maxim) is morally irrelevant.[20]

Note that an asymmetry arises here regarding the assessment of maxims and act tokens. Maxims are assessed directly by the FUL; act tokens, by way of contrast, are assessed indirectly on the basis of the maxim on which they are performed. Thus, for example, in the *Metaphysics of Morals* Kant says that 'you

---

[19] Note that it is still open to me to take on board Kant's distinction between maxims and rules as set out in the *Critique of Practical Reason*. That is, an agent might adopt one maxim on the basis of another one, in which case the former maxim is 'subordinate' to the other. But an agent also then might adopt a technical rule about how to act on the subordinate maxim (and as a technical rule rather than a maxim, the principle would not be assessed by the FUL). Moreover, there always will be room for judgment regarding how to apply a maxim to a specific situation, and great sensitivity must be used in trying to determine an agent's maxims through time given that our values tend to evolve continuously as we grow.

[20] It is worth emphasizing that, strictly speaking, the FUL is *not* a test of maxims: it is a test of act tokens, and it assesses these act tokens on the basis of their corresponding maxims.

must thus [i.e., according to the FUL] first observe your actions *according to their subjective principles*' (MM, 6: 225.08-09, my emphasis). What this means is that when the act token in question is the adoption of another maxim, the im/permissibility of the maxim adopted is therefore independent of the im/permissibility of the act of adoption itself.

My account can be contrasted with Herman's. According to Herman, the permissibility of a maxim of nonbeneficence is determined by examining the permissibility of the maxim underlying its adoption: 'Knowing what is behind a maxim's adoption is...essential to its assessment by the CI procedure' (Herman 1993: 65). I think that this is mistaken: I think that Herman here is conflating the assessment of act tokens with the assessment of maxims.

If I am right about this, it helps to explain a more fundamental confusion of act types and act tokens that underlies the Mixed Maxims Objection. Any given act type can be performed on multiple maxims, and the act token that manifests any given maxim instantiates multiple act types. Moreover, the deontic status of an act type need not be the same as the deontic status of the associated act token or maxim. Thus although putting on warm clothes might be a permissible act type, some instances of putting on warm clothes (such as when those clothes are loaded with a magnet that will trigger a nuclear warhead) might be impermissible. If the egoist's act token of putting on warm clothes really is an instance of acting on the egoist maxim, then (I contend) the egoist's act *token* is wrong even though the type is permissible. And that helps to explain Parfit's confusion regarding this case.

However, my solution to the Mixed Maxims Objection does not explain how the category of the obligatory can be used. To do that, I want to build on the following claim from the *Metaphysics of Morals*: 'Duties toward [*gegen*] one's fellow humans from the respect proper to them are expressed only negatively, that is, this duty of virtue is expressed only negatively (through the prohibition of the contrary)' (MM, 6: 464.32-465.02). I want to draw on what Kant says here to set out my own positive solution concerning how to apply the category of the obligatory using the FUL.[21] Doing so also will help me to explain how the FUL can be used to assess act types.

The argument I want to make is: (a) given what we know about human nature and (b) given what we know about the circumstances in which we live, it may be inferred (using the FUL) that (c) an agent who does not adopt a maxim that involves helping others will have adopted an impermissible maxim. If this argument works, it does not show that any particular act token of helping is obligatory, nor does it show that there is a positive duty to adopt any particular maxim of helping. Rather, it shows that generally speaking, the action type 'helping' is obligatory.

Some might find it surprising that I would appeal to empirical knowledge of human nature in this context. So I think it is worth pointing out that Kant appeals to facts about human nature to derive duties throughout the *Metaphysics of Morals*. Indeed, Kant tells us to expect as much in the general introduction:

> ...we will often have to take the special *nature* of humans, which is cognized only through experience, as an object in order to *show* in it the consequences from universal moral principles, without that however thereby taking away anything from the purity of the latter, or its *a priori* origin being made thereby doubtful. —That is as much as to say: a metaphysics of morals cannot be grounded on anthropology, but nevertheless [it can] be applied to it. (MM, 6: 217.01-08)

I think there are a number of facts about human nature that are relevant for present purposes. One is that we have, as Kant himself remarks, various 'true needs', and these needs span various domains (MM, 6: 393.30).[22] They include the most basic needs for food, water, and sleep, and they include more complex needs that are layered on top of these, like social recognition, understanding, and self-realization. Obviously these needs are more or less pressing in different people, and there is some rough, perhaps Maslovian, hierarchy of importance regarding them. But the point is that most everyone has these needs,

---

[21] In this passage Kant is contrasting duties of respect with duties of love. He thus seems to be suggesting that duties of love might not be expressed 'only negatively'. So the following, insofar as it is meant to apply to all duties, involves some departure from the letter of the text.

[22] See (Rawls 2000: 173-175) or (Kahn 2019: 253n22).

and most everyone knows it. And this leads me to the second relevant fact about human nature: we are aware of the needs of others.

No doubt some people are more sensitive (and some are more oblivious) than others. But for most of us, if we see someone in distress, we will recognize that they are in distress. The recognition might come from subconscious processing of facial cues; it might come from verbal or visual signals; or it might come from simple recognition of an individual's circumstances. But the point is that generally speaking it is not so hard to tell when others are in a bad way. Moreover, Kant seems to think that some sort of capacity for moral sympathy, which he calls love of one's neighbor, is a presupposition of moral agency as one of the '*subjective* conditions of the receptivity for the concept of duty' (MM, 6: 399.08-09).

The third fact that is relevant for present purposes is that our thought patterns are relatively limited and relatively similar. This might be a result of the fact that our hardware for generating thought is relatively limited and relatively similar. It also might be a result of conditioning, or conditioning in limited and relatively similar environments. But that is not the point. The point is simply that the infinitely many thoughts and intentions that might enter a person's mind when ze sees and reacts to another in need are not so variegated as to defy comprehension or, in fact, categorization.

Turning now to the environment (premise (b)), I want to point out that Kant appeals to this kind of consideration, too, throughout the *Metaphysics of Morals*. For example, in his discussion of the impermissibility of suicide, Kant leaves open whether it is permissible to kill oneself in special circumstances, as when someone throws themself into battle 'in order to save the fatherland' (MM, 6: 423.19) or when someone 'feels hydrophobia already, as the effect of the bite of a mad dog' (MM, 6: 423.32-33).

The main fact about the environment that is relevant for current purposes is that we live in relatively harsh conditions. For one thing, we live in conditions of relative scarcity. Even the richest among us will have needs they cannot fulfill on their own. And modern economies leave most in circumstances in which hard work is required to satisfy even their most basic needs at a subsistence level. The situations of many, for social and/or non-social reasons, are quite precarious and known to be so.

One reason why this is known to be so is that land is among the resources that are scarce. As Kant points out, we live on the surface of a relatively small spheroid; 'if [the Earth] were an infinite plane, humans could disperse themselves thereon so that they came into no community with one another at all' (MM, 6: 262.23-25). The Earth is not an infinite plane, and we are bound to bump up against one another as we go through life. More than that, social structures in modern states make it inconceivable that someone could pass through life without being confronted with the fact that others, often others with whom the person will have no actual human contact, are in dire need of help of one form or another: it is discussed in school, plastered on billboards, broadcast on the radio, advertised on television, the stuff of the internet, and just generally 'in your face'.

Given all of these facts, I suggest that if an agent goes through life without performing any helping actions, she will have adopted an impermissible maxim. The impermissible maxim might be one of indifference toward others, it might be an egoistic maxim, or it might be a maxim of harm. But the point is that the maxim will be impermissible, and thus performing some helping actions in one's life is obligatory (because never doing so is indicative of viciousness).

Now let me try to make this argumentative move (from (a) and (b) to (c)) clearer. I am moving from an absence of act tokens to the adoption of an impermissible maxim, and I am making this move on the basis of facts about humans and the human condition (facts about our minds, our environments, our social conditions, and the conditions of those around us). I am *not* inferring the impermissibility of a specific maxim from the absence of any act tokens of a given type. Rather, I am inferring the presence of a maxim that can be determined impermissible by the FUL on the basis of the absence of any act tokens of a given type: I am inferring the presence of a nonuniversalizable maxim on the basis of (i) (near) universal facts about humans and the human condition and (ii) (near) universal facts about how that maxim would manifest in action given (i).

There are two things that are worth emphasizing about this argument. First, it is based on generalizations and, thus, the conclusion itself is general. There might be a realistic life-scenario in which an agent never helps anyone else but never adopts an impermissible maxim. I find such a scenario hard to imagine. But the point of this argument is not to show that such a scenario is impossible or that there is a

strictly universal duty to help others. Rather, the goal of this argument is to show that there is a general duty to help others and, thus, that the act type helping is obligatory.

Second and as noted above, this argument does not show that any particular act token or maxim is obligatory. It builds on some of the ideas in the arguments from section 3 (like the idea that the conditions that require an agent to engage in deliberation about helping others are pervasive). And it attempts to capture the intuitions that those arguments were attempting to capture (like the intuition that helping others is a duty). But it does not attempt to go as far as those arguments.

On my account, the FUL can be used to assess the im/permissibility of maxims directly. The im/permissibility of an act token can be determined indirectly: it corresponds to that of the maxim it instantiates. Precisely because (I) obligatoriness of maxims cannot be determined using the FUL and (II) the deontic status of act tokens corresponds to that of their maxims, it follows that (III) obligatoriness of act tokens cannot be determined.

Nonetheless, if my argument works, then the obligatoriness of an act type can be determined by the FUL indirectly using facts about the im/permissibility of maxims and general knowledge about us. And I want to suggest that the obligatoriness of act types is all that is needed: given the range of maxims and act tokens available in a given instance, to say that one is obligatory seems counterintuitive. When we say, for example, that people should pay their debts, we are talking about the act type, not an act token, and in general there is a great deal of permissibility with regard to how agents pay their debts (with cash, with a check, with a credit card; with the bills in a stack, fanned out, folded; etc.).

I also want to suggest that this is the most charitable way of understanding the nature of the duties Kant discusses in the *Metaphysics of Morals*. Kant's arguments for these duties are based on generalizations about us and about the conditions in which we live, and I think that Kant's inclusion of casuistical questions, which seem genuinely intended to be open for discussion, illustrates his awareness that precisely because the duties are based on generalizations, the duties themselves are general.

I want to close out this section with two brief remarks on how the foregoing general duty derivation fits into other current discussions in this problem space.

First, consider the underdemandingness objection to Kantian accounts of beneficence. This objection is the counterpoint to the overdemandingness objection to utilitarianism, an objection that gained prominence in the wake of Peter Singer's 1972 'Famine, Affluence, and Morality'.[23] The worry for Kantianism is that it will not require someone to aid a drowning child because at least some Kantian accounts of the general duty of beneficence allow for exceptions when an agent is not inclined to be beneficent.

Paul Formosa and Martin Sticker recently have responded to this objection. They balance the general duty of beneficence against the general duty not to be indifferent to the plight of others. They conclude that although a committed humanitarian with a long history of helping others permissibly could refrain from saving a drowning child, '[a]bsent such a record of helping others, failure to help in drowning child cases amounts to vicious indifference and is therefore wrong' (Formosa and Sticker 2018: 636).

The argument I have developed here, however, does not support such a response. On the one side, absent some plausible generalizations about the kinds of maxims that a committed humanitarian might adopt in such a case, there are no grounds for thinking that the humanitarian permissibly may ignore a drowning child. On the other side, absent some plausible generalizations about the kinds of maxims other kinds of agents would adopt in such a case, there are no grounds for thinking that a non-humanitarian may not permissibly ignore a drowning child. Moreover, such generalizations afford at most deliberative presumptions on my account: the deontic status of the behavior of an agent in any given case can be determined only by assessing that agent's actual maxims.

Second, consider recent debates about the difference between perfect duties, like the duty not to lie, and imperfect duties, like the duty of beneficence. For example, Mavis Biss maintains that whereas perfect duties manifest as prohibitions, imperfect duties require a kind of positive moral striving that manifests very differently: 'we can easily imagine an agent who takes moral prohibitions very seriously, resisting temptations to commit vicious acts, and yet ignores or avoids positive moral striving' (Biss 2018: 611).

---

[23] I take this account of the underdemandingness objection from (Formosa and Sticker 2018: section 2).

Biss might be right about the im/perfect duty distinction in general. However I think it is worth pointing out that my argument for the obligatoriness of helping appeals only to ideas about the universalizability tests insofar as they make various maxims impermissible. So the general duty of beneficence that I have derived does not play into a negative/positive distinction of the kind that Biss is making—and deliberately so given what was said in section 3 above.

Conclusion

In this paper I have attempted to do three things. First, I argued that recent replies to Parfit's Mixed Maxims Objection fail, in part because they make categories like the impermissible impossible to apply, in part because they overlook important deontic categories like the obligatory. Second, I argued that recent and independent attempts to show how to apply the category of the obligatory to maxims on the basis of the FUL fail, in part because they presuppose illicit conceptions of maxims, in part because they have unacceptable results. Third and finally, I provided my own solution to these problems, drawing on remarks Kant makes in the *Metaphysics of Morals*.

In closing I want to point out that there is much more work to be done. In addition to the permissible, the impermissible, and the obligatory, there are categories (hinted at above in section 2) like the suberogatory and the supererogatory. Traditionally Kant's ethics has been taken to founder on the supererogatory, and some philosophers recently have questioned whether the FUL is conceptually rich enough to capture intuitions about moral 'ordering', like the idea that killing is worse than stealing. Some of these ideas are present in Kant's work, especially the *Metaphysics of Morals*, which has a general categorization of our duties and includes general remarks about moral merit. But they are beyond the scope of the present investigation. One thing, however, that tells in favor of the approach I have been trying to develop here is that I think it points the way toward how to solve at least some of these other problems.[24]

Bibliography

Anscombe, Elizabeth. (1957) *Intention*. Oxford University Press.

Biss, Mavis. (2018) 'Positive Morality and the Realization of Freedom in Kant's Moral Philosophy'. *European Journal of Philosophy,* 27, 610-24.

Formosa, Paul and Sticker, Martin. (2018) 'Kant and the Demandingness of the Virtue of Beneficence'. *European Journal of Philosophy,* 27, 625-42.

Freyenhagen, Fabian. (2012) 'The Empty Formalism Objection Revisited'. In Thom Brooks (ed.), *Hegel's Philosophy of Right* (Wiley-Blackwell), pp. 43-72.

Gressis, Rob. (2010a) 'Recent Work on Kantian Maxims I'. *Philosophy Compass*, 5.3, 216-27.

Gressis, Rob. (2010b) 'Recent Work on Kantian Maxims II'. *Philosophy Compass*, 5.3, 228-39.

Guyer, Paul. (2006) *Kant*. Routledge.

Guyer, Paul. (2007) *Kant's* Groundwork for the Metaphysics of Morals. Continuum Press.

Harrison, Jonathan. (1957) 'Kant's Examples of the First Formulation of the Categorical Imperative'. *Philosophical Quarterly*, 7.26, 50-62.

Herman, Barbara. (1993) *The Practice of Moral Judgment*. Harvard University Press.

Herman, Barbara. (2007) *Moral Literacy*. Harvard University Press.

Hernandez, Jill. (2010) 'Impermissibility and Kantian Moral Worth'. *Ethical Theory and Moral Practice*, 13, 403-19.

Kahn, Samuel. (2019) *Kant, Ought Implies Can, the Principle of Alternate Possibilities, and Happiness*. Lexington Press.

Korsgaard, Christine. (1996) *Creating the Kingdom of Ends*. Cambridge University Press.

Nyholm, Sven. (2015) 'Kant's Formula of Universal Law Revisited'. *Metaphilosophy*, 46.2, 280-99.

O'Neill, Onora. (2013) *Acting on Principle*. Cambridge University Press.

Parfit, Derek. (2011) *On What Matters Volume 1*. Oxford University Press.

Pogge, Thomas. (2004) 'Parfit On What's Wrong'. *The Harvard Review of Philosophy*, XII.1, 52-9.

Rawls, John. (2000) *Lectures on the History of Moral Philosophy*. Harvard University Press.

Singer, Peter. (1972) 'Famine, Affluence, and Morality'. *Philosophy and Public Affairs*, 1.1, 229-43.

Stratton-Lake, Philip. (2000) *Kant, Duty, and Moral Worth*. Routledge.

Timmermann, Jens. (2000) 'Kant's Puzzling Ethics of Maxims'. *The Harvard Review of Philosophy*, VIII, 43-52.