Why positive duties cannot be derived from Kant's Formula of Universal Law

Ever since Hegel famously objected to Kant's universalization formulations of the Categorical Imperative on the grounds that they are nothing but an empty formalism, there has been continual debate about whether he was right. In this paper I argue that Hegel got things at least half-right: I argue that even if negative duties (duties to omit actions or not to adopt maxims) can be derived from the universalization formulations, positive duties (duties to commit actions or to adopt maxims) cannot.

The paper is divided into three main sections. In the first, I set out the procedures generally accepted among Kantians for deriving positive duties from the universalization formulations. In the second, I set out the arguments from section 1 in more detail and explain why they do not work. In the third, I examine a strategy that might be used to supplement the arguments from section 2 and I argue that it also does not work.

Section 1. How to derive a positive duty

Kant introduces two versions of the universalization formulations of the Categorical Imperative (CI) in the *Groundwork to a metaphysics of morals*. The first is introduced in part I:

> I ought never to proceed otherwise than so, that I also could will, my maxim ought to become a universal law.[1]

The second is introduced in part II:

> Act so, as if the maxim of your action should become through your will a universal law of nature.[2]

There are a number differences, superficial and non-superficial, between these formulations. Nonetheless, Kant clearly takes them to be equivalent at least in the sense that they will yield the same deontic status for any maxim, and I do not propose to argue otherwise.

Kant explains that there are two different ways a maxim can fail to be universalizable:

> Some actions are so performed [*beschaffen*] that their maxim cannot once be thought without contradiction as a universal law of nature; far from it that one yet could will it should become one such. With others indeed this inner impossibility is not to be met, but it is nevertheless impossible to will that their maxim be raised to the universality of a law of nature because such a will would contradict itself.[3]

Kant uses these two ways a maxim can fail the universalization tests to distinguish "strict or narrower (unremitting)" duties from "wider (meritorious)" duties, a distinction he evidently takes to be equivalent to the

---

[1] GMS, AA 04: 402.07-09, emphasis omitted.

[2] GMS, AA 04: 421.18-20, emphasis omitted.

[3] GMS, AA 04: 424.03-10, emphasis omitted. Note that, strictly speaking, these remarks apply only to the second formulation, the law of nature formulation of the CI.

distinction between perfect and imperfect duty.[4] The first kind of contradiction is usually referred to as a Contradiction in Conception (CC), the second as a Contradiction in Willing (CW).[5]

In order to illustrate how these tests are supposed to generate positive duties, it will be useful to have two maxims to work with. I am going to use a lying-promise maxim and a non-benevolence maxim:

> **Lying-Promise Maxim:** I will tell a lying promise in order to get some ready money.
> **Non-Benevolence Maxim:** I will never help anyone.

One reason for using these two maxims is that, as will be seen below, they are frequently used as test cases in the literature with which I am engaging.[6] A second is that they illustrate that the positive/negative duty distinction cuts across the CC/CW distinction (and, thus, the perfect/imperfect distinction). That is, the lying-promise and non-benevolence maxims are both supposed to generate positive duties (to tell only truthful promises and to help others, respectively), yet the lying-promise maxim fails the CC test whereas the non-benevolence maxim fails the CW test.

In order to explain why these maxims cannot be universalized, we need the idea of a maxim's universalized typified counterpart.[7] A maxim's universalized typified counterpart is the law of nature analog of the original maxim and, thus, it expresses causal necessity. The universalized typified counterparts of the lying-promise and non-benevolence maxims are thus:

> **Lying-Promise Maxim-U:** Everybody will tell lying promises in order to get some ready money.
> **Non-Benevolence Maxim-U:** Nobody will help anyone.

There are two main interpretations of how a contradiction is supposed to manifest between a maxim and its universalized typified counterpart and thus two main interpretations of how a maxim fails the CC and the CW tests. These are the Logical Contradiction Interpretation and the Practical Contradiction Interpretation.[8]

According to the Logical Contradiction Interpretation of the CC test, a maxim fails the CC test if (but only if) the action in the maxim would not be possible in a world in which the universalized typified counterpart of the maxim is a law of nature. Thus, for example, in a world in which Lying-Promise Maxim-U is a law of nature, the convention of promise-keeping would not exist. From this it follows that making a promise would be impossible in such a world and (*a fortiori*) a lying promise would be impossible, whence it may be seen that the lying-promise maxim fails the Logical Contradiction Interpretation of the CC test: it would be impossible to conceive of an agent acting on the lying-promise maxim in a world in which its universalized typified counterpart (Lying-Promise Maxim-U) is a law of nature.

---

[4] The quotations in this sentence are taken from GMS, AA 04: 424.10-11. For Kant's remarks about perfect and imperfect duty, see GMS, AA 04: 421n. The context makes the intended equivalence clear (Kant also seems to take this distinction to be equivalent to another one, the distinction between owed duties and duties of love (mentioned at GMS, AA 04: 430n)).

[5] Following (O'Neill, 2013, originally published as Nell, 1975; for the CC test see 143*ff*; for the CW test see 173*ff*).

[6] These two maxims are, as readers of Kant immediately will recognize, taken from the second and fourth examples he uses to illustrate how the law of nature formulation is supposed to work (GMS, AA 04: 429.15-430.27). The first and third examples involve duties to oneself rather than duties to others (the duty not to commit suicide from self-love and the duty to promote one's natural talents).

[7] As with the CC/CW terminology, I believe that the universalized typified counterpart terminology first was introduced by (O'Neill, 2013, 141*ff*).

[8] Ever since (Korsgaard, 1985; reprinted as chapter 3 of Korsgaard, 1996) it has been conventional to divide interpretations of the universalization formulations into three main kinds: the Logical Contradiction Interpretation, the Practical Contradiction Interpretation, and the Teleological Contradiction Interpretation. I ignore the Teleological Contradiction Interpretation in the main text above because very few commentators seem to think it is philosophically viable (they reject the teleological understanding of laws of nature on which it builds).

According to the Practical Contradiction Interpretation of the CC test, a maxim fails the CC test if (but only if) the end that is to be achieved by means of the action in the maxim, although perhaps still achievable by some other means, no longer can be achieved in the original way embodied in the maxim in a world in which its universalized typified counterpart is a law of nature. Obviously if the action is impossible it no longer will be a suitable means to any end, much less the end in the original maxim. So if a maxim fails the Logical Contradiction Interpretation of the CC test, it also will fail the Practical Contradiction Interpretation of the CC test. But it does so for a different reason: the lying-promise maxim fails the Practical Contradiction Interpretation of the CC test because lying promises are no longer an effective means to securing ready money, not because lying promises are no longer possible.[9]

There is less consensus regarding the Logical Contradiction Interpretation and Practical Contradiction Interpretation of the CW test than there is regarding the CC test. On one influential understanding of the CW test, it appeals to "true human needs."[10] The idea is that finite, imperfect human agents have various true needs, like food, water, sleep, and social interaction.[11] So a maxim fails the Logical Contradiction Interpretation of the CW test if its universalized typified counterpart would render nonexistent some means necessary to the satisfaction of the agent's true needs, and a maxim fails the Practical Contradiction Interpretation of the CW test if its universalized typified counterpart would render ineffectual some means necessary to the satisfaction of the agent's true needs. As with the lying-promise maxim, the non-benevolence maxim will fail both interpretations of the CW test: humans are such that we require help to attain our true needs and, if the Non-Benevolence Maxim-U were a law of nature, that help would be nonexistent (and, thus, ineffectual).

A massive (and ever-growing) body of literature has emerged of attempts to refine or replace the Logical Contradiction Interpretation and the Practical Contradiction Interpretation.[12] However, it is important to realize that this literature is independent of the thesis advocated in this paper. That is because, in showing whether some maxim is universalizable, the Logical Contradiction Interpretation and Practical Contradiction Interpretation (and other interpretations) show only whether these maxims are permissible. This is not equivalent to showing that some maxim is obligatory. As will be seen shortly, the part of the argument for positive duties that I am going to attack builds up from the conclusion that a given maxim is not universalizable. That is, it uses the results of the CC and CW tests to show that some maxim is obligatory. But which interpretation is used to generate these results is immaterial. It is for this reason that the issues I am engaging with here are independent from the debate about the nature of the contradiction between a maxim and its universalized typified counterpart. To put this another way: none of the contradiction interpretations, in itself, entails positive duties, and none of the derivations of positive duties is specific to one or another of the contradiction interpretations.

So why are positive duties taken to be so important? If an ethical principle is able to show only that there are maxims that ought to be *omitted*, maxims agents should *refrain* from adopting, this is taken to be a deficiency of that principle. Thus, for example, Herman asserts that "positive duties…are needed to complete an account of individual rational agency."[13] Because of this Kantians try to protect the universalizability tests from having the deficiency just mentioned.

---

[9] This opens up the possibility for maxims that fail the Practical Contradiction Interpretation of the CC test but do not fail the Logical Contradiction Interpretation of the CC test.

[10] (Rawls, 2000, 173-175; see also chapter V section 6).

[11] According to Seymour, "our true needs are defined by those goods which are necessary just to find satisfaction in living…[so] we can allow that our true needs include many things including health, financial and physical security, rest, recreation, freedom, privacy, and social relationships" (Seymour, 2007, 91).

[12] For a critical survey of five commentators' recent (all published within the last 10 years) such attempts see (Kahn, 2019).

[13] (Herman, 2007, 255). Along the same lines, Korsgaard asserts that "in the sort of theory I am defending, responsibility in general is going to look a lot more like responsibility for omission" (Korsgaard, 2009, 175). If there is only responsibility for omission (no responsibility for commission) and if there are no positive duties (no duties of commission), then no agent ever will be responsible for failing to fulfill a duty.

To see how this work is done we can return to Kant. Consider the following two passages, the first from the *Groundwork to a metaphysics of morals* and the second from the *Metaphysics of morals*:

> When we now attend to ourselves in this overstepping of duty, then we find that we do not actually will that our maxim should become a universal law, for that is impossible, but rather that *the opposite* of it should rather remain universally a law; only we take to ourselves the freedom, for ourselves, or (even only for this [one] time) to the advantage of our inclination, to make thereof an exception.[14]

> ...according to them [categorical imperatives] certain actions are allowed or disallowed, i.e. morally possible or impossible, some of these however, or *their opposites*, are morally necessary, i.e. obligatory...[15]

Note that in the first passage Kant suggests that when an agent acts on an impermissible maxim, she wills that the *opposite* of that maxim should be a universal law, which seems to be to take adopting the opposite of the maxim, or perhaps its corresponding action, to be obligatory. Similarly, in the second passage Kant suggests that adopting the opposites of some maxims are obligatory. Kant gives a concrete example of this in section 29 of the *Metaphysics of moral*s when, after remarking that there is a duty to take care of oneself and one's body, he claims that depriving oneself of what one needs to enjoy life is the "*opposite*" of this duty.[16]

Many commentators build on these ideas about maxim-opposites to derive positive duties from the universalizability tests:

1. Paton says that "Kant manages to extract from this formula [the universalizability tests] positive duties, like the duty...of helping others...by showing that to will the opposite [the non-benevolence maxim] is to fall into self-contradiction."[17]

2. Korsgaard argues that "[i]t is only if a maxim fails the [universalizability] test that we get a duty—the duty of doing the opposite of what the failed maxim says, or, more precisely, of adopting the opposite of what the maxim says as universal law."[18] She illustrate this with the lying-promise maxim, arguing that agents who will this maxim are "committed to willing the law that people should make their promises in good faith."[19]

---

[14] GMS, AA 04: 424.15-20, my emphasis. See also GMS, AA 04: 420.08-11.

[15] MS, AA 06: 221.25-28, my emphasis.

[16] MS, AA 06: 452.16-22, at 18 my emphasis. Gregor translates *Gegentheil* as "contrary" rather than "opposite" in this passage. I think that this is mistaken (see note 26 and note 27 below).

[17] (Paton, 1947, 142).

[18] (Korsgaard, 1996, 63). Korsgaard repeats this in subsequent chapters of the same book (originally published as separate articles) and in two subsequent books:

> Using this formulation we discover that a maxim must be regarded as a law by discovering that the opposite maxim—the maxim of not doing the action in question—cannot be regarded as a law. (Korsgaard, 2008, 184)

> A maxim is fit to be a law...[in] the sense corresponding to obligation, if it not only can but must be a law. The way we ascertain this is by showing that the maxim of doing the opposite is unfit to be a law, and must be rejected. (Korsgaard, 2009, 16)

[19] (Korsgaard, 1996, 64). See also (Korsgaard, 2009, 16): "when we find that we cannot will the maxim of false promising just to get a little ready cash, we arrive at the duty of promising sincerely for purposes of this kind."

3. Kim suggests that a maxim is morally obligatory "if and only if it is universalizable, but its opposite is not" and, like Korsgaard, illustrates this by appeal to the lying-promise maxim: "For example, this is the case with respect to the maxim of rejecting false promising."[20]

However, Kant does not give a sustained treatment of the concept of a maxim-opposite, and different commentators explain the concept (and how it is to be used to derive positive duties) in different ways. There are currently three main camps in the secondary literature.

Some commentators appeal to negation:

1. Stohr asserts that "[a]n action is required if the negation of its maxim fails the [universalizability] test."[21]
2. Harrison maintains that "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."[22]

The argument in these passages is that if there is a negative duty not to adopt some maxim, M (i.e., if M fails the CC test or the CW test), in which case the corresponding action, A, is impermissible, then the action referred to in M's negation (M's negation being the maxim produced by substituting "not-A" for "A" in M), namely not-A, is obligatory. Hence there is a positive duty to do not-A. I shall call this the negation derivation.

Some commentators appeal to a related but different line of argument, one that builds on the concept of a maxim-contradictory:

1. According to Herman, "if the CI procedure shows that it is impermissible to adopt the maxim, 'to never help anyone [the non-benevolence maxim],' it follows that we must adopt its contradictory, 'to help some others sometimes.'"[23]
2. Guyer argues that "just as the falsity of one statement implies the truth of its contradictory...so it would seem that the impermissibility of adopting one maxim does imply the necessity of adopting its

_____

[20] (Kim, 2015, 160). It should be noted that, in the passage from which the quotations above are excerpted, Kim equates "morally worthy" with "morally obligatory." This equation is idiosyncratic and threatens to complicate matters, so I have left it out of the main text. The full passage is as follows:

> This considerations suggests that one can define the predicates "morally worth" (or "morally obligatory"), "morally unworthy" (or "morally forbidden") and "morally neutral" ("morally permissible") in terms of the universalizability of maxims. A maxim is morally worthy if and only if it is universalizable, but its opposite is not. For example, this is the case with respect to the maxim of rejecting false promising. A maxim is morally unworthy if and only if it is not universalizable, but its opposite is, e.g., stealing someone else's property.
> Finally, a maxim is morally neutral if and only if both it and its opposite are universalizable.

Kant thinks that an action is morally worthy only if it is performed from duty, and (absent faulty beliefs or faulty wiring) an action can be performed from duty only if it is obligatory. So if an action is morally worthy (as Kant uses the term), then it is obligatory. But the converse of this is false: obligatory actions can be performed from motives other than duty. So Kim's equation of these two categories is mistaken.

[21] (Stohr, 2011, 50).

[22] (Harrison, 1957, 52).

[23] (Herman, 1993, 63). She repeats this line of reasoning with regard to the non-benevolence maxim in (Herman, 2007, 208).

contradictory, thus the impermissibility of the maxim of never helping anyone [the non-benevolence maxim] does imply the necessity of the maxim of helping someone sometime."[24]

3. Stratton-Lake asserts that "an action is morally required if its maxim is universalizable while the maxim of the contradictory act is not...a principle is morally required if it is the contradictory of a maxim that cannot be conceived or willed as a universal law."[25]

I shall call this the maxim-contradictory derivation, where a maxim-contradictory differs from the negation of a maxim in that the former negates the quantifier appended to the action (sometimes/never) whereas the latter negates the action itself (A/not-A).[26] Before I explain the concept of a maxim-contradictory and how it fits into the associated derivation in more detail, I want to note two caveats.

First, Harrison also appeals to the concept of a maxim-contradictory. I nonetheless have categorized him as a proponent of the negation derivation rather than the maxim-contradictory derivation. I do so on the basis of his claim that "if it is our duty to do A, then it is our duty to do non-A." As will be seen momentarily, the maxim-contradictory derivation is different.

Second, Stratton-Lake appeals to contradictory acts as well as contradictory maxims. Because of this he might be seen as a bridge between the negation derivation and the maxim-contradictory derivation. Perhaps he would subscribe to both.

As may be seen from the Guyer quotation above, the maxim-contradictory derivation builds on an analogy between proposition contradictories and maxim contradictories. Two propositions are contradictories if but only if the truth of either entails the falsity of the other and *vice versa*. By way of contrast, two maxims M1 and M2 are contradictories if but only if all agents have adopted M1 or M2 and none has adopted both. The similarity between proposition contradictories and maxim contradictories lies in the fact that if an agent has adopted either M1 or M2, this entails that she has not adopted the other (and *vice versa*). Thus, maxim contradictories typically have the form: "sometimes to A" and "never to A." Now if all but one of an agent's options are impermissible, then the only permissible option is obligatory. So if M1 and M2 are contradictories, in which case they exhaust the agent's options, and one is permissible whereas the other is not, then the permissible one is obligatory. In the quotations above, proponents of the maxim-contradictory derivation apply this to the non-benevolence maxim to show that there is a positive duty to adopt its contradictory, the maxim "to help some others sometimes."

The third kind of argument for positive duties appeals to the concept of a maxim-contrary rather than a maxim-contradictory:[27]

---

[24] (Guyer, 2007, 143). However in an earlier publication Guyer adopts what I am going to call the maxim-contrary derivation: "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).

[25] (Stratton-Lake, 2000, 88).

[26] Textual support for this interpretation comes from Kant's parenthetical glosses of opposite (*Gegentheil*) as "*contradictorie*" (for example, MS, AA 06: 384.05-10, at 06: "*contradictorie oppositum*"; and Anth, AA 07: 230.13-19, at 17-18: "*contradictorie s. logice oppositum*").

[27] It is notable that, in addition to speaking of the *Gegentheil* of duty, Kant also speaks of the *Widerspiel* of duty. For example, he says that making the principle of one's own happiness the determining ground of the will is "*Das gerade Widerspiel*" of the principle of morality (KpV, AA 05: 035.07-08, at 07); that "*Das Widerspiel*" of preserving one's animal nature is willful physical death (MS, AA 06: 421.10-14, at 13); that lying is "*das Widerspiel*" of truthfulness (MS, AA 06: 429.04-07, at 06); and that wishing ill upon others and rejoicing when ills befall them is "*das gerade Widerspiel*" of neighborly love (MS, AA 06: 460.07-11, at 10).

Moreover, when Kant glosses *Gegentheil* as "*contradictorie*" (see note 26 above) he glosses *Widerspiel* as "*contrarie.*" So I think the maxim-contrary derivation's Kantian pedigree is secure.

I note in passing that Gregor translates "*Widerspiel*" as "opposite" at MS, AA 06: 460.10, a translation which, in view of the foregoing, I think is misleading (see note 16 and note 26 above).

1. Hernandez asserts that in Kantian ethics an action is obligatory if but only if its "contrary maxim conflicts with the moral law."[28]
2. According to O'Neill if we test a maxim and its contrary and "if only the contrary cannot be universalized, acts conforming to the maxim are obligatory."[29] She illustrates this with various examples including the non-benevolence maxim: "a maxim of giving *no* help to any other in need is contrary to duty…[thus] Duty requires the adoption of the contrary of 1 [the non-benevolence maxim], "I will do some of what is needed to help the needy.""[30]
3. Freyenhagen contends that "[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."[31]

Two maxims are contraries if but only if no agent can adopt both at the same time.[32] Unlike maxim contradictories, however, it is possible for an agent to adopt neither a maxim nor its contrary. This opens up a gap from the maxim-contradictory derivation: proponents of what I am going to call the maxim-contrary derivation cannot appeal to the premise that, by definition of a maxim contradictory, all agents have adopted either one or the other of a maxim-contradictory pair. When it comes to maxim contraries, some agents might have adopted neither. To plug this gap, proponents of the maxim-contrary derivation note that if an agent is deliberating about some course of action, she must adopt *some* maxim.[33] It follows, then, that if the options an agent is deliberating about are contraries (often taken to have the form, as with maxim contradictories, "sometimes to A" and "never to A") and only one is permissible, the permissible one is obligatory.[34]

To summarize: there are three main derivations of positive duties from the universalizability tests, the negation derivation, the maxim-contradictory derivation, and the maxim-contrary derivation. The first builds on the premise that if some action A is impermissible, then it is obligatory to do not-A; the second builds on the premise that if two maxims are contradictories and only one is permissible, then all agents ought to adopt the permissible one; and the third builds on the premise that if two maxims are contraries and only one is permissible, then all agents who are deliberating about the course of action ought to adopt the permissible one. In the next section I shall explain these strategies in more detail and argue that all of them fail.

Section 2. Making Things Precise

I offer the following enumerated reconstruction of the negation derivation:

1. If some action A is impermissible, then it is obligatory to do not-A.
2. If it is obligatory to do not-A, then there is a positive duty to do not-A.
3. Therefore, if some action A is impermissible, then there is a positive duty to do not-A.
4. Some actions are impermissible.
5. Therefore, there are positive duties corresponding to the actions in 4.

---

[28] (Hernandez, 2010, 403).

[29] (O'Neill, 2013, 166).

[30] (O'Neill, 2013, 183).

[31] (Freyenhagen, 2012, 56).

[32] O'Neill defines the concept of a maxim contrary slightly differently. She says that two maxims are contraries if but only if their universalized typified counterparts are contraries (O'Neill, 2013, 162n25).

[33] The possibility of ending someone's deliberation by (e.g.) a blow to the head is not relevant in this context.

[34] Korsgaard's argument leads me to believe that she would fall into this camp. See (Korsgaard 1996, 152-153 or Korsgaard, 2008, 184).

The problem with this argument is that either 1 is true and 2 is false, or 2 is true but 1 is false. Let me explain.

Suppose that "to do not-A" is interpreted as "not to do A." Then 1 is trivially true, trivially because it is true by definition: it is impermissible to A if but only if it is obligatory not to A. But now 2 is false, given Kant's moral theory. What distinguishes a negative duty (a duty not to A) from a positive duty (a duty to A) is that the former requires merely that the agent not adopt the maxim corresponding to the required action whereas a positive duty requires adopting the maxim corresponding to the required action, and it is possible for an agent to fulfill the former by adopting no maxim at all. For example, an agent who is sleeping fulfills the negative duty not to tell a lying promise or adopt the lying-promise maxim without at the same time positively doing anything or adopting any maxim.[35] So if "to do not-A" is interpreted as "not to do A," then 1 is true but 2 is false.

So suppose that "to do not-A" is interpreted differently, as "to do some B where B is not A." Then 2 is true: if it is obligatory for an agent to do some B, then he has a positive duty to B. But now 1 is false; from the mere fact that some action A is impermissible, it does not follow, given Kant's account of obligation, that it is obligatory to perform some other action B. In fact, 1 is false on this interpretation of "to do not-A" for the same exact reason that 2 is false on the previous interpretation. From this it may be seen that the negation derivation will work only if there is some set of maxims such that only one maxim in that set is permissible and such that all agents adopt at least one maxim from that set. That is, if there is such a set of maxims, then we could say that the permissible maxim is the negation of the other maxims and, therefore, that when we negate an impermissible maxim, we get an obligatory one. This is where the maxim-contradictory derivation comes in.

I offer the following enumerated reconstruction of the maxim-contradictory derivation:

1. For any set of maxim contradictories, all agents must adopt at least one maxim in that set and no agent can adopt more than one.
2. If only one maxim in a set is permissible and if an agent must adopt one maxim from that set, then that one maxim is obligatory.
3. If a maxim is obligatory, then an agent has a positive duty to adopt that maxim.
4. Therefore, if only one maxim in a set of maxim contradictories is permissible, all agents have a positive duty to adopt that maxim.
5. There are sets of maxim contradictories such that only one maxim in that set is permissible.
6. Therefore, agents have positive duties corresponding to the maxim contradictories in 5.

Before I explain why I think this argument fails, I want to make three clarificatory points.

First, as described in the previous section, in the positive duties literature maxim contradictories are conceived of as coming in pairs (using propositions as a model), most commonly "sometimes to A" and "never to A." The idea is that either an agent would A (e.g., help others) or not; if so, he has the maxim "sometimes to A" (sometimes to help some others); if not, he has the maxim "never to A" (the non-benevolence maxim). I have generalized this reasoning in my reconstruction to allow for sets of any cardinality.

Second, note that the "must" in this reconstruction is one of logical necessity. This builds on the definition of a maxim contradictory. To put this another way, "must" should not be interpreted as moral necessity: the maxim-contradictory derivation does not rest on a question-begging assumption that there is a positive duty to adopt one maxim from a set of maxim contradictories.

Third and finally, consider again the example deployed against the negation derivation: an agent who is sleeping and, thus, not actively infringing the prohibition against the lying-promise maxim because he is not actively doing anything. The maxim-contradictory derivation is able to sidestep this example because agents' maxims do not melt away when they go to sleep. An agent's maxims might include life-governing policies and long-term plans as well as more localized rules of behavior. So although a sleeping agent might not be *doing* anything,

---

[35] I adapt this example from a similar one in (Yaffe, 1999).

he still might have adopted the contradictories of the lying-promise maxim and the non-benevolence maxim respectively and, thus, be fulfilling his positive duties with regard to these maxims.

Nevertheless, the maxim-contradictory derivation does not work. The problem, I think, lies in premise 5: I do not think that there are any such sets. I do not think that there are any such sets because I think that the concept of a maxim contradictory is empty. For any maxim, M, and its supposed contradictory, it is logically possible for an agent to adopt neither of those maxims, and so the supposed contradictory is not an authentic contradictory of M. To see why, note that the truth of the proposition, "agent A never/sometimes would B," does not entail that agent A has the maxim "never/sometimes to B." For instance, although agents alive today do not, and never will, torture baby dinosaurs, agents alive today have not adopted, as a subjective principle of volition, the maxim "never to torture baby dinosaurs." For agents alive today, torturing baby dinosaurs is not on the radar. Similarly, although the ancient Egyptian pharaohs never ate meat grown on confined animal feeding operations, they did not adopt the maxim "never to eat meat grown on confined animal feeding operations." Indeed, they lacked the concepts (that of a confined animal feeding operation, for instance) necessary for doing so.

The maxim-contradictory derivation conceives of maxims as laws that govern an agent's behavior externally to the agent in much the same way that Kant thought that the laws of nature govern bodies in space. But this is to misunderstand the nature of a maxim. A central doctrine of Kant's ethics, perhaps *the* central doctrine, is that rational agents are set apart from non-rational objects by virtue of the fact that they can be governed by the *representation* of a law, and that is what a maxim is supposed to be: an agent's subjective principle of volition.

When the dialectic is shifted to the lying-promise maxim and the non-benevolence maxim, the point is harder to see but nonetheless true: the maxim-contradictory derivation rests on a *very* strong claim, that it is *logically* necessary that agents adopt the lying-promise maxim or the non-benevolence maxim or their contradictories, and yet there is nothing logically necessary about it. It is anyway conceivable for there to be a society without a promise-keeping convention, and any agent living in such a society presumably would not have adopted either the lying-promise maxim or its contradictory.

Note the point being made here: I am not challenging the idea that there is a negative duty to refrain from the lying-promise maxim. Even in a society in which there is no promise-keeping convention, it might be impermissible for an agent to adopt the lying-promise maxim. For example, if you or I were to parachute into such a society, nothing would change about the deontic status of this maxim. The point being made here is prior to this; the point being made here is that it is logically possible for an agent to adopt neither the lying-promise maxim nor its contradictory, whence it follows that these two maxims are not genuine contradictories after all and, thus, the argument of the maxim-contradictory derivation collapses.

This, however, points the way to a counter-objection. The point I made in the previous paragraph might be conceded, but if so, that might only be because it is quite weak. My objection gains a foothold against the maxim-contradictory derivation only because the maxim-contradictory derivation rests on such a strong premise. But it would be easy to retool the maxim-contradictory derivation to get around my objection and have positive duties emerge unscathed: weaken the premise I am objecting to so that it says only that any agent engaged in deliberation about the maxims in question must adopt one. This is where the maxim-contrary derivation comes in.

I offer the following enumerated reconstruction of the maxim-contrary derivation:

1. For any set of maxim contraries regarding action A, no agent can adopt more than one maxim in that set and all agents engaged in deliberation about A must adopt at least one.
2. If only one maxim in a set is permissible and if an agent must adopt one maxim from that set, then that one maxim is obligatory.
3. If a maxim is obligatory, then an agent has a positive duty to adopt that maxim.
4. Therefore, if only one maxim in a set of maxim contraries regarding A is permissible, all agents engaged in deliberation about A have a positive duty to adopt that maxim.
5. There are sets of maxim contraries regarding actions agents engage in deliberation about such that only one maxim in that set is permissible.
6. Therefore, agents have positive duties corresponding to the maxim contraries in 5.

I want to make three quick clarificatory points again.

First, in the positive duties literature, maxim contraries, like maxim contradictories, are taken to come in pairs (again: "never to A" and "sometimes to A"). And as with the maxim-contradictory derivation, I have generalized the maxim-contrary derivation argument to allow for sets of maxim contraries that come in any cardinality.

Second, the "must" in the maxim-contrary derivation, like the "must" in the maxim-contradictory derivation, is one of logical necessity. The maxim-contrary derivation avoids the problem associated with the maxim-contradictory derivation by considering only agents engaged in deliberation: that is what makes the difference between conceiving of maxims as external laws (like the laws of nature) and conceiving of them as internal principles.

Third and finally, to bolster these ideas about deliberation, proponents of the maxim-contrary derivation sometimes appeal to what they call the deliberative field. This is a concept that is supposed to capture the intuitive idea that when an agent engages in deliberation, there are some considerations that are salient and that thus loom large in his deliberation, while other considerations are not salient and thus do not factor into the equation. For example, an agent deliberating about whether to adopt the lying-promise maxim does not take into consideration the color of his interlocutor's shirt; that is not a factor that appears in his deliberative field (if it did, he would be engaged in deliberation about a different kind of maxim).

As with the maxim-contradictory derivation, I want to attack premise 5 of the maxim-contrary derivation. But this time my attack will come in two steps. First I am going to argue that maxim contraries often come in sets of cardinality greater than 2. I then use this to motivate the idea that our deliberative options are never constrained in the way required by the maxim-contrary derivation (that is: we are never faced with only one permissible option).

To motivate the first step, think about propositional contraries. Propositional contraries obviously come in sets of cardinality greater than 2. For example:

1. This computer is red all over.
2. This computer is blue all over.
3. This computer is polka dotted all over.
4. This computer is striped all over.

This list could be continued for quite some time. And it is easy to come up with a list of maxim analogs; one has only to think about the kinds of maxims college kids adopt when deciding on a future profession. In fact, thinking about the kinds of maxims adopted in deciding on a future profession leads directly to the second step of my argument because not only are there multiple such maxims from which to choose, maxims that will inform the choices these kids make over long periods of time, but there are multiple *permissible* maxims from which to choose.

Of course, there are also impermissible maxims from which to choose. Aiming to become a professional assassin or drug dealer is presumably impermissible, and there would be negative duties to refrain from adopting maxims associated with these jobs. But that still leaves a large number of permissible professions, no positive duties in sight.

Now a proponent of the maxim-contrary derivation might concede this point but nonetheless not give up on the bigger picture argument. That is, a proponent of the maxim-contrary derivation might argue that although there are no positive duties when it comes to professions, that does not mean there are no positive duties *tout court*. In particular, the considerations I advanced in the previous two paragraphs do nothing to impugn the positive duties to adopt the contraries of the lying-promise maxim or the non-benevolence maxim respectively.

But I think that the kind of deliberation involved in choosing a profession reveals why this objection is a nonstarter. Consider the following two maxims: "I will become a professional philosopher" and "I will become a physician." These two maxims, I contend, represent genuine alternatives about which an agent might deliberate. But as we know, given many of the bioethicists we encounter, these two maxims are not genuine contraries. To get

a list of genuine contraries, we would need an exhaustive list of professions, P1, P2, P3,..., Pn, and then we would have one maxim corresponding to every combination of these. And this set of maxims is not, I think, a set of genuine alternatives about which any agent actually deliberates: not only is the idea of an exhaustive list of professions questionable, but to repeat a point made above, most professions simply will not be on the radar for most agents. Moreover, as large as this set is (it has cardinality 2^n if we include the maxim not to pursue *any* profession on the list) it is incomplete. Someone attracted to both philosophy and medicine might decide to pursue one of these and to leave it open whether to pursue the other. For example, an agent might adopt a maxim to pursue philosophy but to leave for medicine should things go south. Or, more realistically, someone simply might adopt a maxim to pursue one profession and not adopt any principle at all regarding some other.

The line of reasoning explored in the previous paragraph then can be applied to the lying-promise maxim and the non-benevolence maxim. I do not want to challenge the claim that "sometimes to make a lying promise to get some ready money" and "never to tell a lying promise" are genuine contraries. The point I want to make is that these two maxims are not the only options about which an agent might deliberate: an agent also might adopt a maxim "never to tell a lying promise to get some ready money," either because he is unsure about whether he will tell a lying promise in some other conditions (e.g., when a murderer knocks on the door asking for the whereabouts of the intended victim hiding in the basement) or because he simply is not thinking in such grand terms ("sometimes to A" or "never to A"). And this point is even clearer, I think, with regard to the non-benevolence maxim, for the non-benevolence maxim and its contrary are far too abstract for any human agent ever actually to contemplate them as action-guiding principles: "sometimes to help some others" provides no action-guidance, and "never to help anyone" provides too much. In sum, whether we look at the maxims concerning the choice of professions, or instead at the lying-promise maxim or non-beneficence maxim, or instead at maxims of other kinds, I contend that we shall not find a set of maxim contraries regarding actions agents engage in deliberation about such that only one maxim in that set is permissible. This is to say that premise 5 in the maxim-contrary derivation is false.

In the next (and final) section of this paper I am going to consider an objection to this argument. But for now I want to summarize the problems I have raised for each interpretation.

The negation derivation attempts to derive positive duties from the universalizability tests by appeal to the idea that if an agent ought not to A, then he ought to not-A. But as seen above, this line of reasoning does not work: at best it rests on an equivocation on the meaning of the claim that an agent ought to not-A.

The maxim-contradictory derivation attempts to bypass this problem by appealing to maxim contradictories. This opened up a route to positive duties because two maxims are contradictories only if all agents adopt at least one; if a maxim is permissible but its contradictory is not, then it follows that the original maxim is (*eo ipso*) obligatory. The problem with this line of reasoning was that the notion of a maxim contradictory is not well-defined: it treats maxims like laws of nature, principles that govern an agent but that are external to that agent, even though maxims are supposed to be subjective principles of volition: the representations of laws wherewith an agent governs himself.

The maxim-contrary derivation then attempts to bypass this problem by appealing to maxim contraries. Maxim contraries are taken to be subjective principles of volition, and ideas about deliberation are taken to bridge the gap opened up by leaving contradictories behind (agents deliberating about a course of action will adopt one maxim from a set of maxim contraries). The problem with this was that in deliberation, agents do not choose between genuine contraries; they choose from a plurality of different maxims and, contrary to premise 5, in no case is just one of the maxims permissible (whence it follows that no such maxim is obligatory). Now let us look at how the maxim-contrary derivation might be supplemented to avoid this problem.

Section 3. Maxim Entailment

The most frequent objection that has been raised when I present my criticism of the maxim-contrary derivation involves an appeal to ideas about maxim entailment. The appeal goes something like this. Consider two propositions, P1 and P2. If P1 entails P2 and if P1 is true, then P2 is also true. For example, "it never rains" entails

"it never rains on Fridays," so if the former is true then the latter is too. Now consider two maxims, M1 and M2. As may be seen from the lying-promise maxim and the non-benevolence maxim, M1 and M2 are propositions. Thus, there might be entailment relations between M1 and M2 in the same way that there might be entailment relations between any two propositions. This can be used in conjunction with what I call the Maxim Entailment Premise:

> If M1 entails M2 and an agent adopts M1, then she also has adopted M2.

For example, "I will help my child move houses" entails "I will help some others sometimes," so if someone has adopted the former then (by the Maxim Entailment Premise) she has adopted the latter. From this it may be seen that, even if no actual agents deliberate about the contrary of the non-benevolence maxim, that does not entail that no actual agents adopt it. Indeed, if the Maxim Entailment Premise is true, then pretty much everybody does so. This is generally taken to short-circuit my objection to premise 5 of the maxim-contrary derivation. We can reconstruct this argument in detail:

1. If an agent adopts a maxim that entails "never to A," then she has adopted the maxim "never to A." (From the Maxim Entailment Premise)
2. If an agent adopts a maxim that entails "sometimes to A," then she has adopted the maxim "sometimes to A." (From the Maxim Entailment Premise)
3. If "never to A" is impermissible, then it is impermissible to adopt a maxim that entails "never to A." (from 1)
4. If "sometimes to A" is impermissible, then it is impermissible to adopt a maxim that entails "sometimes to A." (from 2)
5. Any maxim about A-ing entails either "never to A" or "sometimes to A." (independent premise)
6. If "never to A" is impermissible, then it is obligatory to adopt a maxim that entails "sometimes to A" (from 3 and 5).
7. If "sometimes to A" is impermissible, then it is obligatory to adopt a maxim that entails "never to A." (from 4 and 5)
8. If "never to A" is impermissible, then "sometimes to A" is obligatory, and if "sometimes to A" is impermissible," then "never to A" is obligatory. (from 1, 2, 6, and 7)

From this it may be seen that the proponent of the maxim-contrary interpretation can skirt my objections to premise 5 using the Maxim Entailment Premise: although agents might not deliberate about sets of contrary maxims with only one permissible option, their deliberative options will entail maxims from such sets, and so if the Maxim Entailment Premise is true, agents will adopt maxims from such sets by entailment.

For example, consider the non-beneficence maxim. When engaged in deliberation about helping others, agents might not deliberate about maxims as broad as "I will never help anyone" or "I will sometimes help others." But any maxim of non-beneficence will entail the former, and any maxim of beneficence will entail the latter. So, because the former is impermissible, it is obligatory to adopt some helping maxim. Thus, it is obligatory to adopt the maxim "I will sometimes help others" by entailment. My attack on premise 5 of the maxim-contrary derivation fails once we take into consideration that an entailed maxim is an adopted maxim.

However, there are at least two problems with this line of reasoning. The first is that the Maxim Entailment Premise is false, leaving premises 1 and 2 unsupported. To see this, note that if the Maxim Entailment Premise were true, it would follow immediately that adopting any maxim would entail adopting infinitely many maxims. For example, if someone adopts the non-benevolence maxim, which entails the maxims "I will never help anyone torture a baby dinosaur," "I will never help anyone or I will become a physician," and infinitely many similar maxims, then it follows from the Maxim Entailment Premise that the person adopts all of those infinitely many maxims. This, I think, is an absurd result, for two reasons. One is that I think it is absurd to ascribe infinitely many maxims to individual human agents. Given our finite capacities, I think that this is absurd in itself, and there

does not seem to be any way to avoid it if disjunctive maxims are possible and if the Maxim Entailment Premise is true.

The other and perhaps more tangible reason why this is an absurd result (which will withstand the rejection of disjunctive maxims) is that there are opacity considerations that arise here every bit as much as they did in consideration of the maxim-contradictory derivation: it makes no sense to ascribe to agents maxims about torturing baby dinosaurs simply on the basis of their adoption of the non-benevolence maxim. Indeed, it is precisely because the Maxim Entailment Premise is *false* on account of opacity considerations that agents often can be persuaded *not* to pursue a course of action by being appraised of what doing so would mean. To put this another way, the problem with the Maxim Entailment Premise is that it involves treating maxims as external laws governing agents' behavior rather than subjective principles of volition.

But as hinted above, there is a second problem with this line of reasoning, and the second problem is independent of the Maxim Entailment Premise. It is that premise 5 is of this new argument is false. Consider the following two maxims: "not to help people torture baby animals" and "not to help people throw babies from rooftops." Neither of these maxims has any entailment relations with "never to help anybody" or "sometimes to help others." Indeed, both of these maxims are permissible, which shows how the problem with premise 5 then undercuts premise 6, which is inferred from it. Moreover, this is not an artifact of starting with a maxim-contrary pair in which the "never to A" version is impermissible. Suppose we start with the lying-promise maxim and its permissible contrary, "never to tell a lying promise." Now consider the permissible maxim "never to tell lying promises to friends," which entails neither the lying-promise maxim nor its contrary. Or consider that an agent simply might not adopt a universal policy about lying promises at the moment. This hearkens back to my argument in the previous section to show that agents do not deliberate about maxim-contrary pairs. But now the point is that agents' deliberative options cannot be mapped onto maxim-contrary pairs by entailment.

Based on this I conclude that the appeal to ideas about maxim entailment does not help: the Maxim Entailment Premise and premises 1, 2, and 5 in the comeback to my criticism of the maxim-contrary derivation are false or unsupported.

Conclusion

In this paper I have argued that positive duties cannot be derived from Kant's universalization tests. In section one, I documented the many interpreters who have tried to derive positive duties from these tests and I categorized their derivations into three kinds: the negation derivation, the maxim-contradictory derivation, and the maxim-contrary derivation. In section two, I explained in detail how these derivations are supposed to work, and I explained where I think they go wrong. In section three, I confronted an objection about maxim adoption by entailment, and I argued that this objection does not work. If the arguments in this paper are successful, then a popular strand of thought about the universalization tests may be laid to rest.

Bibliography

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

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* Vol. 7, No. 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* Vol. 13, No. 4: 403-419.

Kahn, Samuel. 2019. "Defending the Traditional Interpretations of Kant's Formula of a Law of Nature," *Theoria* Vol. 66, Issue 158: 76-102.

Kim, Halla. 2015. *Kant and the Foundations of Morality*. Lexington Press.

Korsgaard, Christine. 1985. "Kant's Formula of Universal Law," *Pacific Philosophical Quarterly* Vol. 66, Issue 1-2: 24-47.

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

Korsgaard, Christine. 2008. *The Constitution of Agency*. Oxford University Press.

Korsgaard, Christine. 2009. *Self Constitution*. Oxford University Press.

Nell, Onora. 1975. *Acting on Principle*. Columbia University Press.

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

Paton, H. J. 1947. *The Categorical Imperative*. Hutchinson's University Library.

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

Seymour, Melissa. 2007. *Duties of love and Kant's doctrine of obligatory ends*. Proquest.

Stohr, Karen. 2011. "Kantian Beneficence and the Problem of Obligatory Aid," *Journal of Moral Philosophy* Vol. 8: 45-67.

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

Yaffe, Gideon. 1999. "'Ought' implies 'can' and the principle of alternate possibilities," *Analysis* Vol. 59, No. 3: 218-22.