4: 400. Reminder. ‘Duty is the necessity of an action from respect for law’. The moral worth of an action doesn’t lie in its effects; it doesn’t lie in aimed-for effects either (4: 401), but in conformity to law alone. What could such a law be? 4: 402. Only the idea of universal law as such. Hence the imperative: I ought never to act except in such a way that I could also will that my maxim should become a universal law. Illustration: the lying promise. The distinction between truth-telling from prudence, and from duty. I cannot at the same time will the lie, and a universal law where all lie. (4:404. The advantage of practical reason over theoretical, and how in the moral realm, anyone is as good as a philosopher.)

I ought never to act except in such a way that I can also will that my maxim should become a universal law...Take this question, for example. May I not, when I am hard pressed, make a promise with the intention of not keeping it? Here I readily distinguish the two senses which the question can have—is it prudent, or is it right, to make a false promise? The first no doubt can often be the case. I do indeed see that it is not enough for me to extricate myself from the present embarrassment by this subterfuge: I have to consider whether from this lie there may not subsequently accrue to me much greater inconvenience than that from which I now escape...To tell the truth for the sake of duty is something entirely different from doing so out of concern for inconvenient results...Suppose I seek, however, to learn in the quickest way and yet unerringly how to solve the problem ‘Does a lying promise accord with duty?’ I have then to ask myself ‘Should I really be content that my maxim (the maxim of getting out of difficulty by a false promise) should hold as universal law (one valid both for myself and others)? And could I really say to myself that every one may make a false promise if he finds himself in a difficulty from which he can extricate himself in no other way?’ I then become aware at once that I can indeed will to lie, but I can by no means will a universal law of lying; for by such a law there could properly be no promises at all, since it would be futile to profess a will for future action to others who could not believe my profession or who, if they did so over-hastily, would pay me back in like coin; and consequently my maxim, as soon as it was made a universal law, would be bound to annul itself. (4: 402)

Section II: Transition from popular moral philosophy to metaphysics of morals

4: 406-9. Kant emphasizes the a priori character of the enterprise, arguing that the derivation of the concept of duty from the common use of practical reason is not at all a derivation from experience, and is compatible with there being no certain empirical examples of actions done from duty. Experience does not teach us what duty is, nor does it teach us which actions have in fact been dutiful ones. The moral law holds for all rational beings as such, and with absolute necessity. The use of examples in ethics must be measured up to the moral law; the moral law is not derived from examples.

The Categorical Imperative

4: 412-4. Laws and imperatives. While everything works in accordance with laws, only rational beings act in accordance with the representation of laws. That is what it is to have a will. Such laws are recognized as objectively necessary, but—given our limitations—subjectively contingent; and therefore perceived as commands or imperatives. The distinction between categorical and hypothetical imperatives, and how the categorical imperative is binding independently of any (or any contingent?) end or purpose.

A hypothetical imperative has the form ‘if you want X, then do Y’. Kant’s categorical imperative has no ‘ifs’ about it. It is absolute (no exceptions for special circumstances), universal (every rational agent in these circumstances should do it), has its source in reason alone (not e.g. in empirical facts about desire, or human nature as sociable etc.), knowable a priori, independent of experience.

4: 415-8. What about happiness? Happiness is an end which can be ascribed surely and a priori to every human being, as a matter of natural necessity. Since it is in this sense a ‘necessary’ end, could it not be a source of categorical imperatives? (Cf. Aristotle) Kant
argues that prudence, i.e. skill in choice of means to happiness, commands only hypothetically. Moreover it is a subjective and contingent matter what this or that man counts as his happiness. One cannot act on determinate principles for the sake of being happy, and no-one can be certain what would make them truly happy; so there are only counsels of prudence. Only morality commands categorically.

4: 419-21. Formula of Universal Law. How are categorical imperatives possible? (Cf. the question of the First Critique: how are synthetic a priori judgements possible?) This question can be fully addressed only in Section III, but a beginning can be made by exploring the mere concept of a categorical imperative: it ‘contains’, beyond the law, only the necessity that the maxim be in conformity with this law. What can that conformity amount to? As Kant has already said in Section I, nothing other than ‘the universality of a law as such’. Hence the single categorical imperative, expressed in terms of the formula of universal law: *act only in accordance with that maxim through which you can at the same time will that it become a universal law.*

4: 422-4. Kant’s illustrations: suicide, lying promise, wasting one’s talents, refusing to help others. The first two, when universalized, involve a *contradiction in conception*: there couldn’t, he says, *be* a law prescribing universal self-killing, or universal lying. The second two involve a *contradiction in the will*: even if there could be *a* law prescribing universal talent-wasting, or universal non-helping, you couldn’t will it.

The End in itself, and the Formula of Humanity

4: 427-8. The will, which ‘determines itself to act in conformity with the representation of certain laws’. The will must have an ‘objective ground’, i.e. it must be directed towards some end. If this end is given by reason alone it must hold for all rational beings. Relative ends can only be the basis of hypothetical imperatives.

Suppose there were something the existence of which in itself has an absolute worth, something which as an end in itself could be a ground of determinate laws; then in it, and in it alone, would lie the ground of a possible categorical imperative, that is, of a practical law.

The thesis that rational nature exists as an end in itself, which is the basis for the categorical imperative, is expressed in terms of the Formula of Humanity: *So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means.*

Kant is supposing there is something that satisfies 3 value conceptions: (i) *End in itself*: has worth unconditionally, independent of desire, valid for all rational beings (Wood calls this an ‘objective end’). Contrast: relative end. (ii) *Existent end* (or ‘self-sufficient end’): something that already exists, and whose existence is in itself an end’, i.e. having worth as something to be esteemed, preserved, furthered. Contrast: end to be effected (zu bewirkender Zweck). The latter is the usual notion of end, but Wood says there is space for the former given the more general conception of end as anything ‘for the sake of which’ we act. (iii) *Absolute worth*: this may mean (i), but may mean ‘dignity’, a value that cannot be compared to or traded off against something else. (See Allen Wood, *Kant’s Ethical Thought*, ch. 4.) These three ideas are conceptually distinct: in principle something could be an end in itself without being an existent end; an end in itself might lack absolute worth; existent ends may fail to be ends in themselves, and lack absolute worth; something with absolute worth might be a relative end and an end to be affected. Kant argues that only humanity is an end in itself, an existent end, having absolute worth.

The Formula of Humanity

4: 429-30. Kant’s illustrations for the Formula of Humanity are the same as for the previous Formula: suicide, the lying promise, wasting one’s talents, refusing to help others. Suicide is not treating humanity in one’s own person as an end in itself; the lying promise uses a person merely as means to an end he cannot share (theft likewise); failure to develop one’s talents, or help others, admittedly does not treat humanity as a mere means: however it fails to positively promote humanity as an end.