Chapter 1: Moving from common-sense knowledge to philosophical knowledge about morality

Nothing in the world - or out of it! - can possibly be conceived that could be called ‘good’ without qualification except a GOOD WILL. Mental talents such as intelligence, wit, and judgment, and temperaments such as courage, resoluteness, and perseverance are doubtless in many ways good and desirable; but they can become extremely bad and harmful if the person’s character isn’t good - i.e. if the will that is to make use of these gifts of nature isn’t good. Similarly with gifts of fortune. Power, riches, honour, even health, and the over-all well-being and contentment with one’s condition that we call ‘happiness’, create pride, often leading to arrogance, if there isn’t a good will to correct their influence on the mind . . . . Not to mention the fact that the sight of someone who shows no sign of a pure and good will and yet enjoys uninterrupted prosperity will never give pleasure to an impartial rational observer. So it seems that without a good will one can’t even be worthy of being happy. Even qualities that are conducive to this good will and can make its work easier have no intrinsic unconditional worth. We rightly hold them in high esteem, but only because we assume them to be accompanied by a good will; so we can’t take them to be absolutely ·or unconditionally· good. •Moderation in emotions and passions, self-control, and calm deliberation not only are good in many ways but seem even to constitute part of the person’s inner worth, and they were indeed unconditionally valued by the ancients. Yet they are very far from being good without qualification - ·good in themselves, good in any circumstances· - for without the principles of a good will they can become extremely bad: ·for example·, a villain’s •coolness makes him far more dangerous and more straightforwardly abominable to us than he would otherwise have seemed. What makes a good will good? It isn’t what it brings about, its usefulness in achieving some intended end. Rather, good will is good because of how it wills - that is, it is good in itself. Taken just in itself it is to be valued incomparably more highly than anything that could be brought about by it in the satisfaction of some preference - or, if you like, the sum total of all preferences! Consider this case: Through bad luck or a miserly endowment from stepmotherly nature, this person’s will has no power at all to accomplish its purpose; not even the greatest effort on his part would enable it to achieve anything it aims at. But he does still have a good will - not as a mere wish but as the summoning of all the means in his power. The good will of this person would sparkle like a jewel all by itself, as something that had its full worth in itself. Its value wouldn’t go up or down depending on how useful or fruitless it was. If it was useful, that would only be the setting ·of the jewel,· so to speak, enabling us to handle it more conveniently in commerce (·a diamond ring is easier to manage than a diamond·) or to get those who don’t know much ·about jewels· to look at it. But the setting doesn’t affect the value ·of the jewel· and doesn’t recommend it the experts.
This ·good· will needn’t be the sole and complete good, but it must be the condition of all others, even of the desire for happiness. So we have to consider two purposes: the unconditional purpose of producing a good will, and the conditional purpose of being happy. The ●former requires the cultivation of reason, which - at least in this life - in many ways limits and can indeed almost eliminate the ●latter goal, namely happiness. This state of affairs is entirely compatible with the wisdom of nature; it doesn’t have nature pursuing its goal clumsily; because reason, recognizing that its highest practical calling is to establish a good will, can by achieving that goal get a contentment of its own kind (the kind that comes from attaining a goal set by reason), even though this gets in the way of things that the person merely prefers. So we have to develop ●the concept of a will that is to be esteemed as good in itself without regard to anything else, ●the concept that always takes first place in judging the total worth of our actions, with everything else depending on it, ●a concept that is already lodged in any natural and sound understanding, and doesn’t need to be taught so much as to be brought to light. In order to develop and unfold it, we shall take the concept of duty, which contains it. The concept of a good will is present in the concept of duty, ●not shining out in all its objective and unconditional glory, but rather- in a manner that brings it under certain subjective ●restrictions and ●hindrances; but ●these are far from concealing it or disguising it, for they rather bring it out by contrast and make it shine forth all the more brightly. ·I shall now look at that contrast·.

My topic is the difference between doing something from duty and doing it for other reasons. In tackling this, I shall set aside without discussion two kinds of case - one for which my question doesn’t arise, a second for which the question arises but is too easy to answer for the case to be interesting or instructive. Following those two, I shall introduce two further kinds of case·. (1) I shan’t discuss actions which - even if they are useful in some way or other - are clearly opposed to duty, because with them the question of doing them from duty doesn’t even arise. (2) I shall also ignore cases where someone does A, which really is in accord with duty, but where what he directly wants isn’t to perform A but to perform B which somehow leads to or involves A. ·For example: he (B) unbolts the door so as to escape from the fire, and in so doing he (A) enables others to escape also. There is no need to spend time on such cases·, because in them it is easy to tell whether an action that is in accord with duty is done ●from duty or rather ●for some selfish purpose. (3) It is far harder to detect that difference when the action the person performs - one that is in accord with duty - is what he directly wanted to do, -rather than being something he did only because it was involved in something else that he directly wanted to do·.  

Take the example of a shop-keeper who charges the same prices for selling his goods to inexperienced customers as for selling them to anyone else. This is in accord with duty. But there is also a prudential and not-duty-based motive that the shop-keeper might have for this course of conduct: when there is a buyers’ market, he may sell as cheaply to children as to others so as not to lose customers. Thus the customer is honestly served, but we can’t infer from this that the shop-keeper has behaved in this way from duty and principles of honesty. His own advantage requires this behaviour, and we can’t assume that in addition he directly wants something for his customers and out of love for them he charges them all the same price. His conduct of his policy on pricing comes neither from duty nor from directly wanting it, but from a selfish purpose. [Kant’s
German really does say first that the shop-keeper isn’t led by a direct want and then that he is. His point seems to be this:- The shop-keeper does want to treat all his customers equitably; his intention is aimed at precisely that fact about his conduct (unlike the case in (2) where the agent enables other people to escape but isn’t aiming at that at all). But the shop-keeper’s intention doesn’t stop there, so to speak; he wants to treat his customers equitably not because of what he wants for them, but because of how he wants them to behave later in his interests. This involves a kind of indirectness, which doesn’t assimilate this case to (2) but does distinguish it from a fourth kind of conduct that still isn’t morally worthy but not because it involves the ‘indirectness’ of (2) or that of (3).]

(4) It is a duty to preserve one’s life, and moreover everyone directly wants to do so. But because of the power of that want, the often anxious care that most men have for their survival has no intrinsic worth, and the maxim Preserve yourself has no moral content. Men preserve their lives according to duty, but not from duty. But now consider this case: Adversities and hopeless sorrow have completely taken away this unfortunate man’s relish for life. But his fate has not made him passively despondent or dejected. He is strong in soul, and is exasperated at how things have gone for him, and would like actively to do something about it. Specifically, he wishes for death. But he preserves his life without loving it, not led by any want or fear, but acting from duty. For this person the maxim Preserve yourself has moral content.

[...] So much for the first proposition of morality: For an action to have genuine moral worth it must be done from duty. The second proposition is: An action that is done from duty doesn’t get its moral value from the purpose that’s to be achieved through it but from the maxim that it involves, giving the reason why the person acts thus. So the action’s moral value doesn’t depend on whether what is aimed at in it is actually achieved, but solely on the principle of the will from which the action is done, irrespective of anything the faculty of desire may be aiming at. From what I have said it is clear that the purposes we may have in acting, and their effects as drivers of the will towards desired ends, can’t give our actions any unconditional value, any moral value.

Well, then, if the action’s moral value isn’t to be found in the will in its relation to its hoped-for effect, where can it be found? The only possible source for it is the principle on which the will acts - and never mind the ends that may be achieved by the action. For the will stands at the crossroads, so to speak, at the intersection between its a priori principle, which is formal, and its a posteriori driver - the contingent desire that acts on it - which is material. In that position it must be determined by something; and if it is done from duty it must be determined by the formal principle of the will, since every material principle - every contingent driver of the will - has been withdrawn from it.

The third proposition - a consequence of the first two - I would express as follows: To have a duty is to be required to act in a certain way out of respect for law.

(1) As for what will result from my action, I can certainly prefer or be drawn to it, but I can’t have respect for it; to earn my respect it would have to be something the will does, not merely something that its doings lead to. (2) Similarly, I can’t respect any want or preference: if the preference is mine, the most I can do is to endorse it; if it is someone else’s I can even love it - i.e. see it as favourable to my interests. What can get respect and can thus serve as a command is something that isn’t (1) a consequence of my volition but only a source for it, and isn’t (2) in the service of my preferences but rather overpowers them or at least prevents them from being considered in the choice I make;
this something is, in a word, law itself. Suppose now that someone acts from duty: the
influence of his preferences can’t have anything to do with this, and so facts about what
he might achieve by his action don’t come into it either; so what is there left that can lead
him to act as he does? If the question means ‘What is there objectively, i.e. distinct from
himself, that determines his will in this case?’ the only possible answer is law. And if the
question concerns what there is in the person that influences his will - i.e. what
subjectively influences it - the answer has to be his respect for this practical law, and thus
his acceptance of the maxim I am to follow this law even if it thwarts all my desires. (A
maxim is a subjective principle of volition. The objective principle is the practical law
itself; it would also be the subjective principle for all rational beings if reason had full
power over the formation of preferences.) So an action’s moral value doesn’t lie in the
effect that is expected from it, or in any principle of action that motivates it because of
this expected effect. All the expected effects - something agreeable for me, or even
happiness for others - could be brought about through other causes and don’t need the
will of a rational being, whereas the highest good - what is unconditionally good - can be
found only in such a will. So this wonderful good, which we call moral goodness, can’t
consist in anything but the thought of law in itself that only a rational being can have -
with the will being moved to act by this thought and not by the hoped-for effect the
action. When the person acts according to this conception, this moral goodness is already
present in him; we don’t have to look for it in the upshot of his action. […]

Chapter 2: Moving from popular moral philosophy to the metaphysic of morals

 […] It is indeed absolutely impossible by means of experience to identify with
complete certainty a single case in which the maxim of an action - however much it
might conform to duty - rested solely on moral grounds and on the person’s thought of
his duty. It sometimes happens that we make a considerable sacrifice in performing some
good action, and can’t find within ourselves, search as we may, anything that could have
the power to motivate this except the moral ground of duty. But this shouldn’t make us
confident that our sense of duty was actually the true determining cause of the will,
rather than a secret impulse of self-love masquerading as the idea of duty. For we like to
give ourselves credit for having a more high-minded motive than we actually have; and
even the strictest examination can never lead us entirely behind the secret action-drivers -
or, rather, behind the pretended action-driver to where the real one secretly lurks -
because when moral worth is in question it is not a matter of visible actions but of their
invisible inner principles. ·

The claim that the concept of duty is an empirical one is not only false but
dangerous. Consider the people who ridicule all morality as a mere phantom of human
imagination overreaching itself through self-conceit: one couldn’t give them anything
they would like better than the concession that the concepts of duty have to come wholly
from experience (for their laziness makes them apt to believe that the same is true of all
other concepts too). This concession would give them a sure triumph. I am willing to
admit - out of sheer generosity! - that most of our actions are in accord with duty; but if
we look more closely at our thoughts and aspirations we keep encountering the beloved
self as what our plans rely on, rather than the stern command of duty with its frequent
calls for self-denial. One needn’t be an enemy of virtue, merely a cool observer who can
distinguish ● even the most intense wish for the good from ● actual good, to wonder
sometimes whether true virtue is to be met with anywhere in the world; especially as one
gets older and one’s power of judgment is made wiser by experience and more acute in
observation. [Kant was 60 years old when he wrote this work.] What, then, can stop us
from completely abandoning our ideas of duty, and preserve in us a well-founded respect
for its law? Only the conviction that
Even if there never were any actions springing from such pure sources, that’s not
the topic. Our concern is not ● with whether this or that was done, but ● with
reason’s commanding - on its own initiative and independently of all appearances
- what ought to be done.

[...] One couldn’t do worse by morality than drawing it from examples. We can’t get our
concept of morality initially from examples, for we can’t judge whether something is fit
to be an example or model of morality unless it has already been judged according
principles of morality. ● This applies even to ● the model that is most frequently appealed
to. Even ● Jesus Christ must be compared with our ideal of moral perfection before he is
recognized as being perfect; indeed, he says of himself ‘Why callest thou me (whom you
see) good? There is none good (the archetype - or model- of good) but one, that is, God
(whom you don’t see)’ [Matthew 19:17; the bits added in parentheses are Kant’s]. But
● don’t think that with God the father we have at last found the example or model from
which we can derive our concept of morality-. Where do we get the concept of God as the
highest good from? Solely from the idea of moral perfection that reason lays out for us a
priiori and which it ties, unbreakably, to the concept of a free will. ● Some have said that
the moral life consists in ‘imitating Christ’, but- imitation has no place in moral matters;
and the only use of examples there is ● for encouragement - i.e. showing beyond question
that what the law commands can be done - and ● for making visible - in particular cases-
what the practical rule expresses more generally. But they can never entitle us to steer
purely by examples, setting aside their true model which lies in reason.

[...] When the thought of an objective principle constrains a will, it is called a
‘command’ (of reason), and its verbal expression is called an ‘imperative’.

All imperatives are expressed with an ‘ought’, which indicates how an objective
law of reason relates to a will that isn’t constituted so as to be necessarily determined by
it - namely, relating to it as a constraint. An imperative says that it would be good to
do or to refrain from doing something, but it addresses this to a will that doesn’t always
do something just because it is represented to it as good to do. Practical good is what
determines the will by means of the thoughts that reason produces - and thus not by
subjective causes but objectively, on grounds that are valid for every rational being just
because it is rational. This contrasts with the thought that it would be nice to act in a
certain way; the latter influences the will only by means of a feeling that has purely
subjective causes, which hold for the senses of this or that person but not as a principle of
reason that holds for everyone.

Objective laws of the good would apply to a perfectly good will just as much to as
to any other; but we shouldn’t think of them as constraining such a will, because it is so
constituted that it can’t be determined to act by anything except the thought of the good.
Thus no imperatives hold for God’s will or for any holy will. The ‘ought’ is out of place
here, for the volition is of itself necessarily at one with the law. Thus, what imperatives
do is just to express the relation of ● objective laws of volition ● in general to the
subjective imperfection of the will of this or that particular rational being - the will of any human, for example.

All imperatives command either hypothetically or categorically. The former expresses the practical necessity of some possible action as a means to achieving something else that one does or might want. An imperative would be categorical if it represented an action as being objectively necessary in itself without regard to any other end.

Since every practical law represents some possible action as good, and thus as necessary for anyone whose conduct is governed by reason, what every imperative does is to specify some action that is necessary according to the principle of a will that has something good about it. If the action would be good only as a means to something else, the imperative is hypothetical; but if the action is thought of as good in itself and hence as necessary in a will that conforms to reason, which it has as its principle, the imperative is categorical.

The imperative thus says of some action I could perform that it would be good, and puts the practical rule into a relationship with my will; and it is no less an imperative if I don’t immediately perform the commanded action simply because it is good - because I don’t know that it is good, and/or because I do know this but my conduct is guided by other maxims that are opposed to the objective principles of practical reason.

A hypothetical imperative merely says that the action is good for some purpose that one could have or that one actually does have. In the former case it is a problematic practical principle, in the latter it is an assertoric one. The categorical imperative, which declares the action to be objectively necessary without referring to any end in view . . . .

After those two kinds of hypothetical imperative [skill and prudence] we come at last to one imperative that commands certain conduct immediately, and not through the condition that some purpose can be achieved through it. This imperative is categorical. It isn’t concerned with what is to result from the conduct, or even with what will happen in the conduct (its matter), but only with the form and the principle from which the conduct follows. What is essentially good in the conduct consists in the frame of mind - the willingness to obey the imperative - no matter what the upshot is. This may be called ‘the imperative of morality’.

If only one could give a definite concept of happiness, then whoever wills the end wills also (necessarily according to reason) the only means to it that are in his power. Unfortunately, however, the concept of happiness is so indefinite that, although each person wishes to attain it, he can never give a definite and self-consistent account of what it is that he wishes and wills - under the heading of ‘wanting happiness’. The reason for this is that all the elements of the concept of happiness are empirical (i.e. must be drawn from experience), whereas the idea of happiness requires the thought of an absolute whole - the thought of a maximum of well-being in my present and in every future condition. Now it is impossible for a finite being - even one who is extremely clear-sighted and capable - to form a definite and detailed concept of what he really wants here on this earth. Consider some of the things people say they aim for! Wealth: but in willing to be wealthy a person may bring down on himself much anxiety, envy, and intrigues. Great knowledge and insight: but that may merely sharpen his eye for the dreadfulness of evils that he can’t avoid though he doesn’t now see them; or it may show
him needs ·that he doesn’t know he has, and· that add to the burden his desires already
place on him. ● Long life: but who can guarantee him that it wouldn’t be a long misery?
● Health: but often enough ill-health has kept him from dissolute excesses that he would
have gone in for if he had been perfectly healthy! In short, he can’t come up with any
principle that could with complete certainty lay down what would make him truly happy;
for that he would need to be omniscient.
So in his pursuit of happiness he can’t be guided by detailed principles but only
by bits of empirical advice (e.g. concerning diet, frugality, courtesy, restraint, etc.) which
experience shows are to be usually conducive to well-being. It follows from this ● that
imperatives of prudence can’t strictly speaking command (i.e. present actions objectively
as practically necessary); ● that they should be understood as advice rather than as
commands of reason; ● that the problem: Settle, for sure and universally, what conduct
will promote the happiness of a rational being is completely unsolvable. There couldn’t
be an imperative that in the strict sense commanded us to do what makes for happiness,
because happiness is an ideal not of reason but of imagination, depending only on
empirical grounds. · This means that whether a person will achieve happiness depends on
countlessly many particular facts about his future states; and there is absolutely no
chance of picking out the actions that will produce the right infinite totality of
consequences that will constitute happiness.

[…] When I have the general thought of a hypothetical imperative, I can’t tell just
from this thought what such an imperative will contain. To know that I have to know
what the condition is. But when I have the thought categorical imperative, I know right
away what it will contain. For all the imperative contains is the law, and the necessity that
the maxim conform to the law; and the law doesn’t contain any condition limiting it
(· comparable with the condition that is always part of a hypothetical imperative ·). So
there is nothing left for the maxim to conform to except the universality of a law as such,
and what the imperative represents as necessary is just precisely that conformity of
maxim to law. So there is only one categorical imperative, and this is it: Act only on that
maxim through which you can at the same time will that it should become a universal
law. Now if all imperatives of duty can be derived from this one imperative as a
principle, we’ll at least be able to show what we understand by the concept of duty, what
the concept means, even if we haven’t yet settled whether so-called ‘duty’ is an empty
concept or not. The universality of law according to which effects occur constitutes what
is properly called nature in the most general sense . . . i.e. the existence of things
considered as determined by universal laws. So the universal imperative of duty can be
expressed as follows: Act as though the maxim of your action were to become, through
your will, a universal law of nature.
I want now to list some duties
1. A man who has been brought by a series of troubles to the point of despair and
of weariness with life still has his reason sufficiently to ask himself: ‘Wouldn’t it be
contrary to my duty to myself to take my own life?’ Now he asks: ‘Could the maxim of
my action · in killing myself · become a universal law of nature?’ Well, here is his maxim:
For love of myself, I make it my principle to cut my life short when prolonging it
threatens to bring more troubles than satisfactions. So the question is whether this
principle of self-love could become a universal law of nature. If it did, that would be a
nature that had a law according to which a single feeling ● created a life-affirming push
and also led to the destruction of life itself; and we can see at a glance that such a
‘nature’ would contradict itself, and so couldn’t be a nature. So the maxim we are
discussing couldn’t be a law of nature, and therefore would be utterly in conflict with the
supreme principle of duty.

2. Another man sees himself being driven by need to borrow money. He realizes
that no-one will lend to him unless he firmly promises to repay it at a certain time, and he
is well aware that he wouldn’t be able to keep such a promise. He is disposed to make
such a promise, but he has enough conscience to ask himself: ‘Isn’t it improper and
opposed to duty to relieve one’s needs in that way?’ If he does decide to make the
promise, the maxim of his action will run like this: When I think I need money, I will
borrow money and promise to repay it, although I know that the repayment won’t ever
happen. Here he is - for the rest of this paragraph - reflecting on this: ‘It may be that
this principle of self-love or of personal advantage would fit nicely into my whole future
welfare, so that there is no prudential case against it. But the question remains: would it
be right? To answer this, I change the demand of self-love into a universal law, and then
put the question like this: If my maxim became a universal law, then how would things
stand? I can see straight off that it could never hold as a universal law of nature, and must
contradict itself. For if you take law saying that anyone who thinks he is in need can
make any promises he likes without intending to keep them, and make it universal - so
that everyone in need does behave in this way - that would make the promise and the
intended purpose of it impossible - no-one would believe what was promised to him but
would only laugh at any such performance as a vain pretence.’

[...] All of the material ends that a rational being voluntarily sets before himself
as things to be achieved through his conduct are merely relative, for their value comes
solely from how they relate to the particular way in which the subject’s faculty of desire
is constituted; and from this we can’t get any practical laws, i.e. any universal and
necessary principles that hold for all rational beings and for every act of the will. So the
only imperatives that these relative ends support are hypothetical ones. But suppose there
were something whose existence in itself had absolute value, something which as an end
in itself could support determinate laws. That would be a basis - indeed the only basis -
for a possible categorical imperative, i.e. of a practical law.

‘There is such a thing! It is a human being! I maintain that man - and in general
every rational being - exists as an end in himself and not merely as a means to be used by
this or that at its discretion. Whenever he acts in ways directed towards himself or
towards other rational beings, a person serves as a means to whatever end his action aims
at; but he must always be regarded as also an end. Things that are preferred have only
conditional value, for if the preferences (and the needs arising from them) didn’t exist,
their object would be worthless. That wouldn’t count against the ‘objects’ in question if
the desires on which they depend did themselves have unconditional value, but they
don’t! If the preferences themselves, as the sources of needs, did have absolute value,
one would want to have them; but that is so far from the case that every rational being
must wish he were altogether free of them. So the value of any objects to be obtained
through our actions is always conditional. Beings whose existence depends not on our
will but on nature, if they are not rational beings, have only relative value as means, and
are therefore called ‘things’; whereas rational beings are called ‘persons’, because their
nature already marks them out as ends in themselves (i.e. as not to be used merely as
means) - which makes such a being ● an object of respect, and ● something that sets limits to what anyone can choose to do. Such beings are not merely subjective ends whose existence as a result of our action has value for us, but are objective ends, i.e. things [Dinge] whose existence is an end in itself. It is indeed an irreplaceable end: you can’t substitute for it something else to which it would be merely a means. If there were no such ends in themselves, nothing of absolute value could be found, and if all value were conditional and thus contingent, no supreme practical principle for reason could be found anywhere.

So if there is to be a supreme practical principle, and a categorical imperative for the human will, it must be one which, being drawn from the conception of something that must be an end for everyone because it is an end in itself, constitutes an objective principle of the will that can serve as a universal law. The basis for this principle is: rational nature exists as an end in itself. Human beings necessarily think of their own existence in this way, which means that the principle holds as a subjective principle of human actions. But every other rational being also thinks of his existence on the same rational ground that holds also for myself; and so it is at the same time an objective principle - one that doesn’t depend on continent facts about this or that subject- - a supreme practical ground from which it must be possible to derive all laws of the will. So here is the practical imperative: Act in such a way as to treat humanity, whether in your own person or in that of anyone else, always as an end and never merely as a means. Let us now see whether this can be carried out. To return to our previous examples: First, someone thinking of committing suicide will, if he is guided by the concept of necessary duty to oneself, ask himself Could my suicide be reconciled with the idea of humanity as an end in itself? - And his answer to this should be No-. If he escapes from his burdensome situation by destroying himself, he is using a person merely as a means to keeping himself in a tolerable condition up to the end of his life. But a man is not a thing [Sache], so he isn’t something to be used merely as a means, and must always be regarded in all his actions as an end in himself. So I can’t dispose of a man by maiming, damaging or killing him - and that includes the case where the man is myself. (This basic principle needs to be refined so as to deal properly with questions such as ‘May I have one of my limbs amputated to save my life?’ and ‘May I expose my life to danger in order to save it?’ I shan’t go into these matters here; they belong to morals and not to the metaphysic of morals.