



Dudley Knowles
**POLITICAL
PHILOSOPHY**

Fundamentals of Philosophy
Series editor: John Shand

**Also available as a printed book
see title verso for ISBN details**

Chapter 2

Utilitarianism

A good way to begin is by studying a deep and well-worked-out ethical theory which has commanded wide assent, reached clear conclusions when tackling the philosophical problems thrown up by our political life and produced unambiguous policy directives to settle practical questions. I select utilitarianism because I believe it has these features (or, at least, makes these claims). This has been recognized by many of the most impressive recent contributors to political philosophy. Few endorse utilitarianism – but most of them see the need to define their position against the utilitarian salient.¹ Utilitarianism should not be treated as a straw target; it has two great virtues which we should not lose sight of. First, it is based on a thought that ought to have universal appeal: when judging conduct, we should pay close attention to the consequences of human actions in respect of their contribution to the welfare of all those whom the actions affect. Second, (and this was a central preoccupation of the classical utilitarian thinkers, Jeremy Bentham and John Stuart Mill) that focus is particularly apt for fixing the purposes of government. We would do well to

recapture the revolutionary impact of the claim that government, in particular, is in business to promote the well-being and reduce the suffering of all of its subjects.²

The foundations of utilitarianism

But first things first. Let me give a summary of the main elements of utilitarian theory, beginning with the simplest formulation of the principle of utility:

Right actions maximize well-being.

This statement can be elaborated in many different ways, although it is worth mentioning now that the most familiar version of the principle, invoking the greatest happiness of the greatest number, should have been abandoned long ago. Recent commentators³ have pointed out that a principle which requires the maximization of two independent variables will be indecisive over a significant range of cases. To use Evans's example, trying to rank outcomes in accordance with the greatest happiness of the greatest number is like offering a prize to the person running the furthest distance in the shortest time! Bentham, who first brought the phrase 'the greatest happiness of the greatest number' to prominence, used it as a kind of standing reminder that everyone affected by policies were to be counted and as a slogan redolent of democratic sentiments, but even he recognized that it was faulty in suggesting that the happiness of the majority only, the greatest number, should be counted. He saw that careless use of the principle in this formulation quickly leads the critic to charge that the utilitarian is prone to ignore the rights of minorities and to countenance other injustices so long as a majority is suited.⁴ As we shall see, these questions cannot be settled quite as quickly as a faulty grasp of the principle suggests. For now let us just repeat that everyone's interests are to count equally in the calculations. As Bentham insisted and Mill repeated: 'Everybody to count for one, nobody for more than one.'⁵

Formal theory

Acts, rules and dispositions

There are two main elements to the utilitarian approach which need to be distinguished and reviewed separately – formal theory and value theory.⁶ Formally, utilitarianism is a consequentialist theory. It requires that we compare alternative outcomes in point of their consequences, asking which realizes the maximum amount of some good. Which good is to be maximized is a matter for value theory and we shall examine candidate goods later. An important first question, then, which my account so far has concealed, concerns *what it is* of which we are to review the consequences. Three answers may be distinguished – and it is an important issue whether or not, or to what degree, they may be combined.

In the first place, and most obviously, we may review *directly* the consequences of alternative actions. The thought here is straightforward: we propose to evaluate the rightness or wrongness of actions by determining what the consequences of actions have been or by projecting what the consequences will be.⁷ This position is often referred to as act utilitarianism. A second proposal is different. The rightness or wrongness of actions should be reckoned *indirectly* in terms of whether or not they are in accordance with moral rules; this is the basis of the rule utilitarian programme, the main burden of which will be the assessment of alternative rules (and systems of rules) to determine which *rules* will promote the best consequences. A different, and perhaps complementary, variety of indirect utilitarianism proposes that we evaluate actions in terms of the motives, dispositions or traits of character they exemplify, for example, praising a person who is honest or criticizing one who is mean. In this case the utilitarian will consider which qualities of character are likely to induce those who exhibit them to act in ways that lead to the maximization of well-being. This is evidently an important question for any moral theory which proposes to address issues concerning the formation of character in processes of moral education. And we would do well to remember that John Stuart Mill believed these issues were central to the utilitarian agenda.

So, we can be act or direct utilitarians, or indirect utilitarians investigating the tendencies of rules or dispositions of character. This way of putting matters suggests that these are alternative routes for the utilitarian to follow, whereas the correct position may be different. Note first of all, that there is no reason in principle why each of these subjects, actions, rules and dispositions, may not be susceptible to utilitarian review. We can examine separately the consequences of actions taken individually, of adopting and maintaining rules of conduct, of producing and promoting dispositions to act in various ways. In fact, any version of utilitarianism worth its salt will need to be able to appraise actions and agents in each of these ways. There will be problems for the utilitarian only if we have reason to think that assessment along these different dimensions will yield contradictions or dilemmas. Unfortunately there is reason to think that it might.

There are good utilitarian reasons for societies to introduce and stick to rules of property, determining who owns what, who may use what and much else. Conflicting claims are reduced, the possibility of co-operation is enhanced. Suppose we have in place a set of rules which will promote the best consequences for everyone if they are generally accepted. They include the familiar commandment: Do not steal. Suppose Sally needs to steal a few potatoes from Robert's field if she and her children are to survive. Robert, we might assume to make the case stronger, has plenty of potatoes to spare; he does not in fact notice the theft – and nor does anyone else. Sally, now well fed, finds work and can support her children. She is never tempted to steal again. Did she do right or did she do wrong? To the rule utilitarian she did wrong. The rule in play prohibits stealing and Sally broke it. The act utilitarian will judge differently. The gain to Sally and her children is evident. Robert's loss is negligible. More good has been achieved by the theft than by the family's starvation. We should conclude that act and rule utilitarianism reach different verdicts in this particular case.

The same conflict of views can be reproduced in cases involving act and disposition utilitarianism. Let us agree that a society which succeeds in creating compassionate and sympathetic dispositions in its members will better promote well-being than one which does not.⁸ Carol gives money to a street collector who uses it to buy arms for a terrorist group. She was credulous in believing

the money would be used to help wives and families in need of support and so contributes to the success of a damaging terrorist campaign. Since dispositions as deep as that inculcated in Carol cannot be switched on and off, her compassion as much as her credulity renders her vulnerable to evil solicitations. The disposition utilitarian will commend her display of compassion. The act utilitarian will say she did wrong if her act resulted in a great deal of suffering. Likewise, in cases where my compassion for others causes me to steal in order to prevent their starvation, the demands of disposition utilitarianism seem to conflict with a utilitarianism of rules.

Does this succession of cases reduce utilitarianism to incoherence – simultaneously condemning and endorsing actions from conflicting stances of judgement? Perhaps not, if we accept the main lines of the following characteristic utilitarian response.

What is the chief impetus behind our insistence that we should take into account the utility of rules and dispositions as well as, directly, the utility of acts? It is this thought: it is fantasy to suppose that the moral agent can be forever computing the respective utilities of all prospective acts in order to judge which is best. We haven't the time, we haven't the patience and, perhaps most important, we haven't the knowledge necessary to reach correct verdicts on what future consequences will follow a host of alternative interventions. This point may seem devastating to the act utilitarian but he has a swift response – which is to insist that if we take into account the utility of deliberating over what we should do we shall soon see that short-cuts are necessary. Why should the sailor start working out when high tide will be at Greenock tomorrow on the basis of what it was on a specific day last month if he can look it up in the *Glasgow Herald* or the *Nautical Almanac*?⁹ Clearly we need some analogue of the tide-tables in morality and moral rules give us one. Instead of working our own way through the likely consequences of our actions, why not refer to a set of rules which provides accurate guidance?

If fallibility and the cost of calculation suggest an important place for rules, they also accord considerable weight to the cultivation of character. Some people do mental arithmetic well – and this disposition can be cultivated – but no one except Jeremy Bentham has suggested that the skills of utilitarian calculation ought to be

widely developed.¹⁰ But many utilitarians (and John Stuart Mill conspicuously) have accepted the importance of inculcating strong dispositions, believing that spontaneous and unreflective responses of generosity and honesty will more than compensate for our defects as calculators.

For the act utilitarian, then, rules and dispositions of character, far from comprising alternative dimensions of utilitarian assessment, must be employed in a practical and reliable calculus. Ultimately, of course, the only measure of right action is the goodness of the consequences of actions but this is not a measure that needs to be taken if we have to hand the right set of rules and a population with correct dispositions.

This account is plausible – but how does it help us tackle the problem caused by the examples of conflict given above? It suggests that we have to understand clearly the relation between actions on the one hand and rules and dispositions of character on the other.

Let us begin with rules.¹¹ Thus far I have been referring to rules as though these are simple phenomena with which we are well acquainted. In fact there are at least three different conceptions of rules in play.

The first sort of rule is the ‘ideal’ rule – a technical device unique to utilitarianism. We introduce ideal rules when we claim that actions are right if they are in accordance with those rules which would promote most good, were they to be generally complied with.¹² Ideal rule utilitarianism has been effectively criticized.¹³ A first difficulty is this: Suppose our car is stuck just below the top of the last hill before we reach our destination. The rule for all five passengers to follow is clearly, ‘Push’, if pushing will see them over the summit and into a comfortable bed. Four passengers either don’t work this out or don’t follow the rule. If ideal rule utilitarianism were the best decision procedure to follow, the fifth passenger should push even if her solitary efforts will prove ineffective. This is clearly irrational. And it doesn’t look like a utilitarian strategy either, since no benefits would accrue and the diligent rule follower will suffer for her efforts.

A second difficulty follows. Alter the scenario so that only the pushing of four passengers is needed to get the car over the summit. Why should the fifth passenger push? Isn’t utility maximized if the fifth passenger loiters alongside rather than lending her

shoulder to the task? Again ideal rule utilitarianism suggests that not pushing would be wrong, although it is hard to see how her unnecessary expenditure of effort could maximize utility. (I accept that other things might be wrong with her not pushing. Perhaps it is unfair of her not to shoulder her share of the burden. But now it looks as though fairness conflicts with utility.) The central point of both these examples is that real utilitarians would not stick to ideal rules if the circumstances dictate that utility is best advanced by breaking them. In J.J.C. Smart's terms, either one is a utilitarian or a rule worshipper – one can't be both.¹⁴

The second conception of rules identifies them as useful rules of thumb. A better example here than Mill's *Nautical Almanac* (which we should treat as infallible!) is a rule for hillwalkers such as 'If you cannot see the point to which you are heading, take a compass bearing and follow it'. Accepting such a rule will lead you to take a map and compass on your walk and generally help you to escape difficulties in route finding. But it is important to recognize that the rule should not be followed slavishly. It should be quickly broken if the bearing takes you to the top of a cliff. And if the compass veers erratically when you seem to be walking in a straight line, you should consider whether there might be magnetic rocks in the vicinity.

Are moral rules like this? There is good reason to think that they are. 'Keep promises', we say, but we recognize lots of cases where exceptions may properly be made. Sometimes we cite a specific qualification to the rule which suggests that the rule is more complex than the original simple formulation. We can add a clause: '... unless the promise has been extorted.' We can gather together exceptions, as when we say: '... unless disproportionate harm will be caused to some third party.' Or we can make exceptions on a case by case basis whenever exigencies seem to require the breaking of the promise. When these things happen, the utilitarian says we are justified if we can maximize well-being by breaking the rule.

It has been argued, successfully I think, that this rules-of-thumb variety of rule utilitarianism reduces to act utilitarianism since the bottom line in each of these calculations is that the right action is the one that maximizes utility. We can expect rules which we employ in the face of uncertainty to develop the force of

inhibitions. We may be taught them in the secure expectation that they will develop this motivational power. But whatever the motivational push or pull exhibited by the rules we endorse, we should not expect them to be either immune to revision or privileged against exception wherever utility dictates. The utilitarian claims, with some justification, that the effects of spontaneous good judgement are so positive that we should be reluctant to break rules without compunction; the purposes served by having rules in the first place will not be easily subverted if the rules are strongly internalized. Thus although this variety of rule utilitarianism is consistent with (because it is reducible to) act utilitarianism, there remain strong reasons for supporting the induction of moral rules like 'Keep promises' in the consciousness of agents – just as there are strong reasons for getting walkers to make a *habit* of using a map and compass.

There is a third conception of rules which is of particular importance in political philosophy. This is the category of rules which are constitutive of institutions. Many of these rules will have the force of law and will be backed by legal sanctions although there are non-legal rules and non-legal sanctions. We can expect most societies to have an institution of private property. Such an institution is best understood in terms of an interrelated set of rules establishing rights, duties, powers and privileges. The core rules will be expressed in law, such as prohibitions against theft. But there will be associated non-legal or non-enforceable rules, too. 'Do not write in books that you borrow from friends' is one which I expect most readers to accept. Other institutions which comprise a mix of legal and non-legal rules include marriage and family life, arrangements for treating the sick and educating the young, and of course the political life of the community.

The 'ontology' of such institutions is complex and is not usefully clarified by the modern fad for issuing mission statements. I take an idealist view. Institutions as I describe them consist in rules which command the behaviour of members, rules with respect to which one may take an internal or external point of view. Internally, members (insiders) identify with institutions whose rules they recognize as valid. The external perspective is taken by observers (outsiders) who describe institutions on the basis of members' conduct. Of course, the same person may be both insider and

outsider; these terms describe roles or perspectives and so all depends on the stance from which he is describing or evaluating the rules in question.¹⁵

Institutional rules differ from rules of thumb in two significant ways. In the first place, they will be justified as necessary for the effective functioning of the institution, serving as means to given ends. This is an oversimple way of describing a matter of great conceptual delicacy since it supposes that the purposes of institutions may be identified independently of the structure of rules which constitute them. But my point is this: suppose we recognize as *one* of the purposes of family life (or of alternative patterns of domesticity) the provision of a healthy and supportive environment for children. We shall then justify rules, both legal and non-legal, in terms of their conduciveness to this purpose.

Now remember that for the moment we are utilitarians. We have institutions characterized by rules which promote whatever purposes the institutions serve. It follows that we do not evaluate institutional rules one by one and directly, in terms of their several contributions towards utility. It will be the institution as a whole which is appraised. The utilitarian will tackle first the grand questions, for example: Should children be brought up in a nuclear family, an extended family or in a kibbutz? Having assessed the respective utilities of these different domestic arrangements, we can then go on to fix e.g. the rules for appropriate income tax allowances or whichever means we employ to support what we have concluded is the optimal domestic unit. Institutional rules differ from rules of thumb in that the primary focus of justification is the institution rather than the rule.

The second major difference is equally important. This concerns the justification of particular actions. Assume that we have in place a system for the regulation of private property which includes rules governing inheritance and bequest. My family are hard up. Am I morally justified in forging alterations to Donald's will so that his estate will give my family the security they deserve rather than support the drug addiction of Donald's intended beneficiary?

If the rules governing bequests were rules of thumb, immediately the question would be open: What does utility dictate in these circumstances? Matters are very different when we are thinking of

institutional rules and it is important to work out just why this is so. Here is one immediate difference. In the case of rules of thumb, the rules have standing in the deliberations of the moral agent as ready reckoners which obviate the need for hard, often fallible, calculation – but where circumstances scream out for judgement outside the normal response of compliance, direct calculation of the appropriate utilities can be the only rational response. In the case of institutional rules, these have an authority quite independent of their service as guides to conduct for the unsure or hard-pressed. They are not open to scrutiny except as elements of institutions which find their justifications in terms of their operation as a whole. One may certainly question an institution, demanding whether or not it promotes utility. But if it does then the institution becomes entrenched, acquiring a social reality which cannot be dissolved by the exercise of deliberation. Similarly, one may seek to alter the institutional rules. Maybe different, better, rules will serve the institution more effectively. And this kind of tinkering goes on all the time, conspicuously in legislative activity. But suppose an institution promotes utility in the way its defenders claim and suppose the rules of the institution effectively secure this. If the utilitarian accepts these claims, it is not open to him to violate the rules in order to promote utility. If two people decide the most worthwhile way to spend their time is by playing chess, so long as the game is proceeding it is not open to one of them to cheat on the grounds that both of them will better enjoy subsequent play. It might indeed be true that it will make for a better game if the rules are changed, and this may prompt them to change the rules, giving a handicap to one of the players. But as the rules stand at the beginning of the game, cheating cannot be vindicated by rule changes it may be sensible to introduce later. The cost of subscribing to institutions which promote utility is that one sacrifices the opportunity of breaking rules on occasions which suggest that rule breaking promises utilitarian gains.

So, if I am caught out in my forgery of Donald's will, I should not expect the officials of the legal system to listen carefully to my utilitarian defence. They will follow the rules which utility has dictated should be followed in all cases. There is no great utilitarian ombudsman prepared to review all instances of individuals claiming they broke the rules in the service of some overriding

utilitarian purpose. Nor should this kind of pleading persuade us that one should be installed. Readers can work out for themselves the disutility of introducing such an institutional role!

This is not to say that in emergencies, in cases of disaster or catastrophe, the rules of institutions should not be broken. You may justifiably break the speed limit driving a badly injured person to hospital. But then you should not expect to get punished either, since an institution which is justifiable and maximally effective will make provision for such cases by, for example, specifying allowable defences against the charge of wrong-doing. If such defences are not in place, then the rules of the institution should be altered to permit them. Contrast this with the rule of thumb about following compass bearings. We don't alter or amend the rule when we find ourselves at the top of the cliff. We disregard it until we have circumvented the obstacle – and we pick it up later. We are not in the business of formulating optimal rules of thumb with guidance for each contingency; such rules would quickly become unwieldy and just as difficult to apply as pristine act utilitarianism. But we are in the business of designing and sustaining optimal institutions and there is something desperately wrong with institutions which cannot be remedied in the face of conspicuous disutility. Commanding officers, we are told, may decorate soldiers for bravery – then punish them, if their heroism involved disobeying orders. 'Change the rules', the utilitarian should insist.

Let us conclude, for the moment, that the utilitarian can endorse two different conceptions of rules: rules of thumb which pre-empt arduous and uncertain calculation and institutional rules which promote utility through the dynamics of some complex system. So rules have a place (or better, two) in utilitarian judgement. Can similar arguments be used to sort out the possible conflict between the utility of actions and the utility of dispositions?

I suspect that they can. It makes sense to cultivate in ourselves and others qualities of character which reduce the possibility of conflict and enhance the prospect of fruitful co-operation. It makes sense to subdue or eliminate tendencies which generate conflict or render it endemic. As utilitarian strategies these look eminently respectable – always supposing that conflict promotes suffering and co-operation increases well-being. Each of us can

draw up our own list of favourite and suspect personal qualities. High on my list of admirable qualities, being in Nietzschean terms a typical modern wimp, would be compassion and sympathy, courtesy to strangers (especially beggars), tolerance and good humour.

One of the nice things about speaking of 'dispositions' in this context is the implication that there are no iron laws dictating rigid uniformities of response. Utility may determine that widespread good temper and equanimity may limit occasions of conflict, but the right thing to do in particular circumstances, again judged in terms of utility, may well be to erupt with rage. Once more, the utilitarian should recommend the cultivation of dispositions to counter the rigours and difficulties of judgement. Common sense tells us that those with a generous temperament are a social asset – but it also tells us that generous responses should be restrained if circumstances suggest that those who look to be in need of assistance would really be better off learning to cope with this kind of difficulty by themselves. In the case of rules of thumb we saw how rules could be of general use even though their application could not be justified in conspicuous cases where utilitarian calculation suggests otherwise. In the case of the utility of dispositions, the whiff of contradiction is more easily dispelled since we have no tendency to think of qualities of character as rigid determinants of action.

One interesting question remains. I suggested in respect of institutional rules that these do have an authority which defies the application of utilitarian calculation to particular opportunities for utility promotion. Is there any analogue with respect to qualities of character? I suspect that there may be. The category of institutions as I have employed it has been very wide,¹⁶ comprising almost any congeries of rules, although I have suggested they will have some structure dictated by function or purpose. In fact, I doubt whether any institution can have the ethical force sufficient to motivate members or subscribers to develop the 'internal' point of view with respect to its rules, if it does not cultivate as qualities of character distinctive and appropriate styles of emotional response. There is something bloodless (and plain wrong) about any analysis of domestic relationships which focuses on rules and concomitant rights and duties as the essence. Capacities for love and affection should be in the foreground.¹⁷ In which case, the

utilitarian claim that specific institutions are conducive to general utility will require that participants display the appropriate emotional qualities – and we should recognize the force these may exert on individuals who exhibit them. Who knows . . . there may be occasions when in defiance of these institutions and their internally necessary sentiments, general utility requires the sacrifice of one's first-born son; unlike God, the utilitarian should not then expect obedience.

Aggregate and maximum average utility

A further question in formal utilitarian theory concerns the matter of whether we are to maximize aggregate or average utility. For most purposes, computation in terms of aggregate or maximum average utility will give an identical ordering of different outcomes. If Policy A produces 100 units of utility and Policy B produces 50 units, Policy A is better on aggregate. If both policies affect the same number of people or apply over the same population, say 50 persons, A will again be better than B because the average of two units per person is greater than the average of one. So long as the number of persons over whom the average is taken is constant between the alternative outcomes, no issue of principle is raised.¹⁸

But this condition will not always hold good. We can all think of policies concerning housing, medicine, pollution control, traffic management even, which *themselves* determine, in part, the number of people affected by the policy. The possibility of population control, government action which is directed towards increasing or, more likely nowadays, decreasing the size of the population of a country, is a particularly obvious example of policy which gives rise to a new theoretical problem for the utilitarian.

Suppose two policies C and D effect the same aggregate utility – 100 units. Should we prefer policy C which distributes these units between 50 people (an average utility of two units per person) or policy D which leads to a doubling of the population and which then distributes the 100 units between 100 people (an average of one unit per person)? In point of aggregate utility the totals are the same; in point of maximum average utility the results are very

different – C is superior to D. Whether the units measure resources or states of mind like happiness, it looks as though we should judge that policy C will make us better off. The cake is divided amongst fewer people. Children work out this principle at a very early age – just as soon as they find that times are harder with every addition to the family. With no more detail to go on, our intuitions favour Policy C.

But we should ask, if we strongly favour being amongst the lucky few, where are the missing 50 people? Do we have nothing to say about them? Have they no claim on us? These questions may strike you as silly, but there is a point to them. We do hypothesize such ‘missing persons’ and consider how policies will impact on them when we think through the consequences of what we do for future generations. I can start thinking now of saving for my grandchildren’s education. I don’t have any grandchildren at the moment and may never turn out to have any, but the idea of planning for these hypothetical descendants is not ridiculous. I must plan for my retirement, or so my independent financial adviser instructs, yet he knows no better than I whether I shall live to enjoy it. It makes sense to think of and plan for persons who do not now exist and may never do so, just as it does for persons who now exist but may turn out not to do so when the plans come to fruition. So, if our choice of policy determines that 50 people who don’t presently exist will never do so, shouldn’t we consider the consequences of what we do for them, what we have deprived them of or spared them from? If we select Policy C rather than Policy D haven’t we denied them the prospect of a life with a positive sum of well-being? And maybe there is a Policy E in the offing which promises 150 units of utility spread between 120 beneficiaries. In this case there is a clear gain in terms of aggregate utility. Isn’t this the best thing to do? This intuition conflicts strongly with the claim that the only thing that matters is maximum average utility. I shall leave this tricky problem unresolved. In what follows I shall be supposing that it is average utility that we are seeking to maximize – but you should note my reservations.

Maximization

One final assumption needs to be openly displayed and this is concealed in the unexamined use of the term ‘maximization’. The standard utilitarian picture is that of agents, in their personal capacity or as policy-makers, charting the consequences of actions and then listing the positive and negative effects as these impact upon individuals. We ‘maximize’ utility by selecting that policy or action, amongst a range of alternatives, which promotes the greatest net utility – and the implication is that we decide on the best option by adding the utility scores in respect of each person affected to produce a sum of utility points represented by a cardinal number for each alternative outcome. Something like this practice was implicit in my discussion of the comparative outcomes in respect of aggregate and maximum average utility in the paragraphs above.

The questions begged by this construal of maximization as addition are many and deep and I cannot begin to explore all their ramifications – but here are a few.¹⁹ Are the good (and evil) consequences of action susceptible of measurement at all? Can the consequences for one person be tallied as the sum of the varieties of ways in which persons may be affected? Suppose a policy both diminishes my liberty and improves my health. On what scale can these different effects find a common measure? If we agree that individuals may be able (somehow) to answer these questions for themselves, how are different individual responses to be compared and then registered in a common scale? To employ the familiar jargon, how are interpersonal comparisons and measurement of utility possible? Two things are absolutely clear: first, that a common denominator amongst a range of goods that will permit the arithmetical operations of addition and subtraction (as well as multiplication and division as soon as probabilities enter the calculations) will be very hard, if not impossible, to find; the utilitarian, for all Bentham’s talk of a ‘felicific calculus’, may well have to manage with instruments of calculation which do not permit the operation of arithmetical functions. Second, just what is possible in both individual and interpersonal cases will depend upon the description of the goods in question – and so it is to utilitarian value theory that I now turn.

Value theory

Utilitarian value theory tells us what to look for when we assess actions, rules or dispositions in the light of their consequences. It tells us what it is that we are measuring when we set out to compare alternative actions or states of affairs and judge which is best. Thus far, I have described the good to be assessed as utility (the weakest formulation), well-being or welfare. I have been supposing that we have a rough idea of what these terms connote, but in truth I have been issuing blank cheques, trusting the reader to fill in the value in a plausible fashion. It is an open question whether the utilitarian has the philosophical assets to redeem them. In this brief survey of different accounts of the value to be maximized I shall highlight issues which have a bearing on the agenda of the political philosopher (although the prime concern of the utilitarian who wishes to contribute to debates in political philosophy will be to give the *correct* account of value!).

Hedonism

The classical utilitarians, Bentham and John Stuart Mill, thought of value, the human good or the good of sentient creatures, as happiness and explained happiness as pleasure and the absence of pain. This identification of the good with happiness is the doctrine of hedonism. For Bentham and Mill it was an empirical claim about human nature that human beings desire happiness – and Mill went so far as to claim that, at bottom, happiness is the *only* thing they desire.²⁰ Mill's strategy in this proof has two elements: happiness is a good, he claims, because everyone desires it, and happiness is the only good because any alternative candidate good can be seen to be either a means to happiness or a part (or ingredient) of it.

Mill's arguments are not easy to evaluate. It is clearly a hedonist position; what is hard to see is whether Mill has successfully disengaged himself from the egoism of Bentham, as he believed. He thought it obviously true that agents desire the happiness of others. They may be kind, helpful, generous and compassionate – and the sensible utilitarian will acquire dispositions of these sorts

and encourage them in others. And having dispositions of these sorts may cause their bearers to act in ways that reduce or sacrifice their own happiness. Mill must insist, at this stage in his proof, that these qualities of character, which we may call virtues, following his account, must be, in some sense, elements of the agent's own happiness. Their life must be going better for the exhibition of them. The virtuous person must be frustrated and diminished if no opportunities arise to be virtuous, since their virtue is a part of their happiness. Minimally we must suppose that the virtuous person enjoys the exercise of virtue, but if we are to steer clear of egoism (and retain some sense that the agent is acting truly virtuously) we must detach the enjoyment from the motivation of the agent.

In fact, this is quite easy to do. I guarantee (unless the circumstances are somehow peculiar) that you will gain pleasure, sometimes great pleasure, from acting virtuously. Many recipes for attaining pleasure are unsound given the ways human beings differ. The sources of pleasure are a matter of self-discovery rather than expert tuition. Nonetheless, the satisfaction of having done something worthwhile is as universal a phenomenon as any that may be attested in this area. And yet it is clear that those who achieve it do not act in order to gain it. It is not a satisfaction that can be actively sought, a sensation that one can pursue with forethought and diligence – and without hypocrisy. It is a very special kind of moral creep (a kind I have not encountered in either real life or fiction) that will react with pleasure at the prospect of someone desperately in need of assistance. 'Oh good!' such a one might exclaim, 'another opportunity to gain that unique kind of satisfaction which I experience when helping others!' I hesitate to generalize over the peculiar sphere of human motivation, but I don't see how the sense that one has acted virtuously can co-exist with the knowledge that one has sought and attained a kind of personal pleasure which one prizes. The fact is that although the feeling of pleasure is just about universally *consequent upon* the genuine exercise of virtuous sentiments, the actions themselves will not be *motivated by* the prospect of attaining it.

Mill knew this very well. But if it is true, what remains of his claim that actions performed by agents who desire to act virtuously are explained by or exemplify the desire for happiness? Of

course virtue can contribute to happiness – I have argued that it always will in the sense of producing in the agent an invariable sense of satisfaction in having acted well – but this is not the same as claiming that virtue is a part or ingredient of happiness. (Cheerfulness and a feeling of content that one’s life is going well: these are the sorts of states of mind that can be recognized as ingredients of happiness.) I conclude that the virtues only contribute to our happiness when it is not our happiness that we seek by their exercise, and hence that acting virtuously is something that we desire independently of the prospects for our happiness, however good these prospects might be as a result. If this is accepted, it follows that happiness is not the only good we seek. We also recognize the good of a virtuous life.

We now have two distinct goods – happiness and the pursuit of a virtuous life. Perhaps there are more. The standard way of developing a list of distinct goods is to give examples of conflict. A famous example, discussed by Griffin,²¹ is that of Freud who was prepared to suffer a great deal of pain during his terminal illness in order to continue being able to think clearly. So one could claim on this basis that knowledge of one’s affairs and one’s surroundings is a good independent of the absence of pain. Amend the example slightly and we can describe cases where a suffering patient is prepared to undergo great pain in order to retain control over those aspects of her life which she judges to be important. This will introduce a separate value of autonomy.²² A slightly different patient may undergo great suffering in order to accomplish some task which has been central to her aspirations – designing a house, planting a forest or writing a book. And we may applaud those who risk their lives climbing mountains, diving caves, undertaking arduous sea voyages – all in the grip of ambitions which cannot be described as the pursuit of pleasure. So it looks as though a sense of achievement is a characteristic human good. Health, too, is distinct from pleasure (and these other goods). I may sacrifice my health in the pursuit of pleasure – and give up pleasurable activities if these threaten my health. I may risk my eye-sight painting miniatures and keep up skiing at the cost of further damage to my knees. The list of distinctive human goods throws up dilemmas at every turn.

In the face of these difficulties the utilitarian may continue to

insist on hedonism, but it is quite clear that he can do so only by continuing to insist that all of these separate goods are desired as the means to happiness or as constitutive of it – parts or ingredients, in Mill’s terms. I suspect however that whatever cogency the argument can gather is achieved by stipulation because our concept of happiness is so ragged. Defined as pleasure and the absence of pain, the concept is operational but, as I suggested above, we are forced to recognize other conflicting goods. If we are to include these competing goods in the account we give of happiness, then happiness becomes little more than a cipher, collecting together all of the distinctive objects of human desire. What threatens, of course, is incoherence since happiness is no longer a value in terms of which we can appraise alternative outcomes which promote happiness along these different dimensions. We have lost the sense of happiness as a common denominator which can be employed in the calculation of what is the best thing to do.

Desire-satisfaction

The utilitarian is still not without resort. He can claim, still with an eye on Mill’s proof, that we have overlooked one important unifying feature, that these goods are each of them the objects of characteristic human desires. In which case, why not identify the satisfaction of desire as the distinctive good to be employed in evaluating outcomes?

To many this has seemed a very attractive proposal. Desires (or preferences) are revealed in human actions. Our actions serve as the mark of their strength; indeed the prices of goods, determined by how much we are prepared to pay, may quantify their intensity and register the degree of our satisfaction. At this point ethics and political philosophy join hands with economics and all the powerful mathematical tools of that discipline are liberated for application outside the conventional boundaries of the dismal science. No longer will we have to pretend to be ‘weighing’ the pleasures and pains in prospect as though these could be put on the scales with fruit and vegetables. Welfare economics is at the disposal of the consumer with a spreadsheet who wishes to take a voyage of self-discovery, as well as being the resource of the policy-maker

concerned to implement those policies which maximally suit those affected by them. Bentham's antiquated apparatus of the 'felicific calculus', computing the intensity, duration, propinquity, fecundity, etc. of pleasures and pains can be consigned to the same museum of primitive scientific instruments which houses the first slide-rule.

It is difficult to overestimate the importance in policy-making of this line of development of utilitarian theory, although the harms caused by its application as well as the dangers in prospect may be considerable. Environmentalists rail at the application of the techniques of cost-benefit analysis to questions involving the conservation of wild nature or beautiful countryside. How can these goods be weighed in the balance?²³ At the moment, however, we are considering its theoretical underpinnings – and these are not secure. There are two initial difficulties which both point in the same direction. In the first place it is obvious that desire-satisfaction may not be a good where the desire is ill-informed or ill-judged. A sick child who hates the taste of medicine may have her strongest desires satisfied when she pours it down the sink, but if the child is ignorant of the properties of the compound or judges that its taste is of greater importance than its curative effects, this preference should be discounted. Its satisfaction is not a good. So we modify the account, seeking value now in *informed* desire satisfaction. Other desires should be subject to scrutiny as well – and this leads us to the second major difficulty. Take the desires of the sadist. It looks as though our evaluation of sadistic behaviour will require us to give some weight to the satisfaction of his desires, with the utilitarian registering these in the balance with the desire of the victim to avoid the pain the sadist is keen to inflict. A squeamish desire-satisfaction utilitarian must hope and pray that the dissatisfaction of the victim is greater than the satisfaction in prospect for the sadist. But surely the satisfactions of the sadist should not count *at all* in the evaluation of his conduct. That his preferences are satisfied when he succeeds counts towards the evil rather than the good of what he does. So again the account needs to be amended; the good to be registered is now the satisfaction of desires which are both fully informed and legitimate; illegitimate as well as ignorant and poorly judged desires should be discounted.

The difficulties are obvious. How much knowledge and rational capacity do we need for our desires to count as sufficiently well-informed? We need more than the child who believes that nothing which tastes awful can do her good – but do we need as much knowledge as the best science makes available before our desires are sufficiently well-informed? How much good judgement do we require, supposing all relevant information is to hand? Again, we shall need more than the child who believes the avoidance of nausea is a greater priority than good health. But how much more is not easy to determine. Smoking, one is told, reduces life expectancy by five years on average. Is there something defective in the judgement of the well-informed doctor who continues to smoke despite the risk to her health?

The response to the sadist example is even trickier. Defects of knowledge and judgement subvert the natural authority of the desires they generate and so there is more than a whiff of normativity in the requirement that desires be well-informed and soundly judged. There must be, in prospect if not in place, canons for the appraisal of desires from these perspectives. And these canons cannot derive from considerations of utility upon pain of circularity in the account. This difficulty is even more evident in the case of the requirement that desire-satisfaction be gained legitimately, since the utilitarian needs a non-moral argument to show that the desire for another's harm, and the satisfaction gained from achieving it, should be entirely discounted.²⁴ The most dangerous tack here would be to distinguish as legitimate desires which are normal or natural, alluding to some spurious hybrid of folk biology and religious dogma, of the kind that powerful churchmen are prone to sell.

I do not believe that the utilitarian has the philosophical and anthropological resources necessary to breathe life into the claim that the fulfilment of desire is the root of all human value or that desirability is the basis of a formal account of the good which collects together all the qualities of life which humans value. If we can describe separately, and vindicate as plausible, a range of human goods, I see no point in adopting a theoretical apparatus which collects them together under one label – as desirables or as ingredients of happiness – if that apparatus does no work in the ranking of outcomes as better or worse. In some cases we may