21

John Rawls

21.1 Rawls' theory of justice

The most complete argument for a theory of justice is possibly that provided by Rawls, who argues for his two principles of justice in *A Theory of Justice* (1972). His theory is of justice as fairness, accepting those principles that would result from an 'original position'. In this 'original position', the parties set out, subject to conditions considered reasonable and fair, to agree the principles by which their society should be organised. It is thus a social contract position, although the contract is a hypothetical one.

Method

Rawls accepts Hart's distinction between concepts of justice and conceptions of justice. He agrees that any theory of justice must deal with both of these. By a concept of justice Rawls means the role of its principles in assigning rights and duties and in defining the appropriate division of social advantages. This is essentially an objective phenomenon. By a conception of justice he means the interpretation of the role of these principles in particular situations. He acknowledges that this is much more subjective.

Rawls' theory in its own terms is designed to cope with situations where mutually disinterested persons put forward conflicting claims to a division of goods and services under conditions of moderate scarcity. His theory is of no application in conditions of total scarcity, for example, Mozambique.

His method is to test all the previous theories and from their defects to extrapolate a superior theory. While he states that his is the best (not surprisingly) he acknowledges the existence of other theories. His starting point is a rejection of utilitarianism.

Rejection of utilitarianism

The first half of the book is directed at a rejection of the earlier theories. We shall concentrate on his comments with regard to utilitarianism. Classical (total) utility is

easily dismissed, since it supports an increase in population even if average utility would be thereby decreased. Average utility is more of a problem; Rawls has criticisms of it. Some of these are general, in terms of allowing sacrifice of persons for others: we have already discussed utility in this light. Some criticisms are specific to the original position (OP) and show why parties to the original position (POP) would not choose average utility. The veil of ignorance denies them knowledge of the probabilities of being at any particular economic/social level in the society. To choose average utility is to take a risk on whether or not one's position is a good one; it is the logical choice if the principle of 'insufficient reason' is followed, and the probabilities for each position are regarded as equal. Rawls says that the POP would reject such risk-taking and would, as we shall see below, adopt a 'maximin' strategy, that is, one that maximises good only so far as a minimum of good is maintained.

Rawls states that utilitarianism suffers from two major defects:

Utilitarianism ignores the distinctness of persons. The utilitarians assume that just as rationality requires making small sacrifices for larger gains, so it also requires a trade-off of the welfare of some against the welfare of others. This idea of trading off the welfare of some against the welfare of others conflicts with our moral intuition. In Simmonds' book *Law, Justice and Rights* the author uses the analogy of having a toothache and going to the dentist and suffering great pain for a few moments in order to have the toothache cured. This is a utilitarian calculation and it is submitted it is a legitimate utilitarian calculation because one is talking about sacrifices for oneself. One is undergoing the pain in order to gain the pleasure for oneself. As indicated in the previous Chapter the problem with utilitarianism is that it speaks of sacrificing the pain of others for the pleasure of oneself.

Furthermore utilitarianism merely regards people as mechanisms to measure pain and pleasure and as receptacles in which welfare is to be maximised with the greatest possible efficiency.

Utilitarianism seeks to define the right in terms of the good. Utilitarianism begins with an account of good and defines as right that which brings about the good. Herein is the anomaly, because utilitarianism takes account of unjustly obtained happiness. According to Rawls,

> 'Justice is the first virtue of social institutions in the same way as truth is the first virtue of thought. And like truth, justice is uncompromising ...'

Characteristics of a theory of justice

Having rejected utilitarianism, Rawls goes on to consider the conditions which any theory of justice must satisfy before it is a useful theory of justice. These are:

1. The theory must be general.
2. The theory must be universal in application.

3. The theory is public in the sense that the population must know about it, there being no point in having a secret theory of justice.
4. The theory must impose an ordering on competing claims.
5. The theory must have finality.

Content

According to Rawls, justice is prior to happiness. It is only when we know that happiness is just that we regard happiness as having any positive value. Is Rawls then affording us a neutral conception of justice? Rawls believes that justice represents the framework within which different individuals have a fair opportunity to pursue their own goals and values. His theory seeks to meet the criticisms that he levelled against utilitarianism and attempts to employ the criteria of rational prudence in a manner consistent with both the distinctness of persons and the priority of the right over the good.

In his book, Rawls uses a complex mixture of two forms of reasoning: deductive and inductive; and of two bases for argument: the contract and reflective equilibrium. We will look at the two bases in turn.

Contract

As we have said, Rawls envisages an original position in which, in conditions of equality (discussed below), the parties agree to the principles to judge society. A contract model is used for two reasons:

Exposition. As a nice framework within which to explain the various conditions which can reasonably be imposed on such an agreement 'the conditions we accept as "fair"' and to show how the principles can be reached from these conditions.

Justification. The contract is not supposed to have been made *in fact*, but the device is used to show 'to emphasise' that the principles chosen are 'fair', ones we would accept given a fair starting point. Anybody at any time can enter the original position and if they did and were rational they would arrive at the same conclusion, the same principles of justice, that Rawls arrives at. For this reason these principles are said to be objective and would be binding on the members of the society. He has thus, it is submitted, overcome the difficulty with the earlier social contractarian theories that could not account for how the contract was binding on those not party thereto. The contract model is thus justificatory, in support of the principles and giving them some degree of legitimacy, as well as expository.

Having established the conditions of the contract, the argument is both deductive and inductive. Rawls attempts to show deductively that the principles would in fact be chosen by the parties, and why they would be chosen. They would value liberty very highly, as the first principle chosen shows: they would then adopt maximin as in the second principle. The two principles result from the conditions of the original position and that deduction.

The main argument, as Rawls calls it, is of another kind. It involves looking at a list of other conceptions of justice, particularly variants of utilitarianism, and seeing why the two principles should be preferred. This argument is not deductive. Nor is it a complete argument: there may be further conceptions of justice not yet disclosed which would win, and anyway Rawls' list is not complete re present conceptions.

In our discussion of Rawls, we concentrate on the deductive argument from the original position.

Reflective equilibrium

How are the conditions in the original position decided upon? Basically, they are conditions we accept as reasonable to impose on (people choosing our) standards of justice; for example the veil of ignorance encourages impartiality. There is, however, another important aspect, that of trying to reach reflective equilibrium.

Rawls outlines the derivation of principles, following a process whereby reasonable conditions are used, and the principles that result are discovered. These principles are tested against our considered judgments in particular situations. If they run counter we must either, after thought, reject our considered judgments, or alter the conditions of the original position. The results of those altered conditions emerge, are checked against our considered judgments, and so on. Eventually, by modifications at both ends, reflective equilibrium emerges between the principles emerging and our considered judgments. We have then reached the best version of the original position, which expresses reasonable conditions and yields principles matching our considered judgments. The starting point, the conditions of the original position, is tinkered with until correct principles emerge.

This argument is clearly not deductive. It also makes the original position itself redundant, if one is being harsh. Why not just put up possible principles and modify them with regard to our considered judgments, without referring to the contract idea at all? We must refer to the earlier points (contract as exposition and justification) to see why Rawls uses a contract even though it is not strictly a necessary part of the argument.

The original position

The parties

Rawls does tell us who the parties will be; they are various representatives, all of the same generation. We will not consider who they are in detail: it is pointless since one or all come to the same conclusion. The conditions are meant to be such that the two principles are the 'correct answer', to which everyone will agree.

What choice?

The parties are to choose general principles by which society should be organised, principles of justice for that society. It is interesting throughout to consider how various other possible theories are eliminated by Rawls. We discuss this below.

As pointed out above, the choice is made primarily by arguing against other conceptions, but also by a positive deductive argument for the various facets of the two principles.

What motivation?

The parties agree to principles that we term 'principles of justice', but it is important to note that from the point of view of the parties the principles are not principles of justice, but the terms of the contract/agreement made on the basis of self-interest in an original position. As we've seen, the conditions of the original position put everyone in the same situation, and 'self-interest' is therefore non-specific; but given the conditions of ignorance, etc, in the original position, the parties are trying to get the 'best deal' for themselves.

How can they do this, bearing in mind that they knew nothing about themselves? The answer is that they will try to get the highest possible total of primary basic goods. One thing that is removed from the parties is an awareness of their own, particular, plan of life, their conception of what is a 'good life', and what, specifically, they need to live it (time, money, power, responsibility, etc). They do, however, know general facts; one of these general facts is that to fulfil a plan of life everyone wants more and not less of the 'primary basic goods'. These goods 'rights, powers, health, etc' fall into two categories, social and natural; the parties will try to arrange liberties, opportunities, powers, self-respect, income and wealth. Everybody's plan of life will be enhanced by these.

The conditions

What conditions are the parties under?

1. Veil of ignorance: the most obviously necessary condition is one to ensure impartiality, that is, that one's own position and views shouldn't influence the choice. This is a traditional feature of justice theories; a similar device to Rawls' is the impartial spectator.

   Rawls' device is a veil of ignorance, behind which the parties are working. The parties do not know their place in society, their status or class, their fortune, level of natural ability, intelligence and so on. All are in a condition of equality: hence 'justice as fairness', and hence unanimity of result (important, according to Rawls, because it shows a genuine reconciliation of interests). The parties would be rational, free and have knowledge of the general situation but no specific knowledge of the particular. They would know that there is a society; they would know there is intelligence; they would know that there are sexes but they would not know where they as individuals would fit in society. Behind this veil of ignorance any knowledge of all those features which distinguish one person from another will be excluded. Rawls argues that it would be just to impose these conditions and that any decision reached in this condition would become binding. The veil would then be lifted to the extent of the proposition and there would be

no possibility of repeal of the principle. Let us see how this works by an example at the end of this section.

2. Non-altruism: the parties must look to their own self-interest and not to the interests of others, or incoherence will result.
3. Non-envy: In any non-equal situation, the problem of envy will arise. If the parties thought they would be envious of anyone who did better than they did, strict equality would be the only choice, and that might mean that increases in wealth for all are missed. To avoid this, Rawls says the parties will not be envious: they will want the best possible result for themselves, but it will not be a factor against an arrangement that someone else might do better.

Great inequalities could be sanctioned in this way: but Rawls' two principles avoid this inequality. The first principle gives liberty to all equally. The second principle does allow the unequal distribution of primary social goods, but within strict limits. In any case, one of these goods being distributed is self-respect, and one's self-respect would clearly be harmed by any massive inequality resulting in a small share of the other goods.

The just savings principle/family feeling: the parties are all from the same generation. What is to stop them from deciding as follows: we cannot alter how previous generations have acted? Motivated by self-interest (non-altruism) we care not for future generations. Therefore we will use up as many of the resources of the earth as we wish, and we will not, individually or as a society, be concerned about what investment we make for future generations.

Logical, yes; just, no. To avoid this conclusion, Rawls added a condition that the parties were concerned about their families including other generations thereof. (This, of course, results in the just savings principle in the second principle.)

Risk aversion: the parties will choose the least worst alternative and not the best possible alternative because of their self-interest that dictates that they avert risk of a worst possible scenario.

Let us look at a worked example of how this process is meant to occur. We have entered the original position. The proposition that is put to us is that there should be accorded more rights to men than to women, in other words that there should be discrimination against women in the allocation of benefits in the society. This is one of the questions a society has to address. This is the question. The parties do not know whether they are men or women. Acting in their own self interest (non-altruism) and seeking to avert risk, the parties will apply the maximin principle seeking to maximise benefit while minimising burden. Let us say that the parties opt for sex discrimination; the consequences are as follows:

1. If they are men then that will be the best possible scenario. .
2. If they are women then that will be the worst possible scenario.
3. If the the parties act according to the maximin principle they will choose that there should be no sex discrimination in which case whether they are men or women will make no difference. True, they will lose the chance to have all the

benefits that accrue to men if there was discrimination, but they will also avert the risk of all the loss were they women.

Thus it seems that the use of this original position together with the conditions stipulated by Rawls will lead to a denial of all arbitrary discrimination in accordance with established liberal thought.

The principles of justice

From the original position, we arrive at the two principles of justice:

1. Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system for all.
2. Social and economic inequalities are to be arranged so that they are both:
   a) to the greatest benefit of the least advantaged consistent with the just savings principle, and
   b) attached to offices open to all in conditions of fair equality of opportunity.

Generally, the first principle is prior to the second.

We will look at the two principles in three stages. First, the deductive argument for the two principles. Second, the 'main argument', rejecting other conceptions. Third, the argument for priority of the first principle over the second, the priority of liberty.

Deductive argument for the two principles

Assumptions. The conditions of the original position are above, but we should mention at this stage that there are two more assumptions, which shape the form and content of the two principles. The first of these is *social co-operation*. No one can succeed in the plan of life without society, therefore everyone will be willing to enter into social co-operation. Talents are pooled to the benefit of all; the major question is how the results of that pooling should be distributed: what distribution should be decided upon?

It is this assumption that knocks out natural rights theories like those of Nozick's. Is social co-operation a successful argument against the millionaire who argues against re-distributed taxation? Rawls would say that without society that million would not be made or increased. Is the millionaire right to say that he gives more (jobs, investment) than he takes out?

The second assumption is *risk aversion*. The veil of ignorance hides from the parties both the position they will occupy in society and the probabilities relating to the various positions. If there are five million poor people out of 500 million, the odds are one in a hundred that any given individual will be a poor person: but the parties don't know this. In this situation the logicians would suggest the acceptance of an 'insufficient reason' standpoint whereby all possibilities are taken as equally likely. Rawls rejects this, and says the parties would instead choose a maximin position, arranging society so that the position of the least well off class is maximised. The parties are then risk averse: rather than chancing a possibly very

poor position by going for an arrangement which produces very good results at the top end, they will choose that situation where the bottom end is as high as possible.

Do you think the parties would be risk averse? Might they decide to gamble?

The argument

The parties set out, then, to derive principles upon which they can agree to arrange society. They wish these principles to result in the best possible arrangement of primary social goods, and they will decide to adopt a maximin outlook.

The starting point is equality: the parties would start by thinking what the principle to be chosen should be, simply that everyone gets equal shares of everything. They would then realise that some inequalities will benefit everyone: incentives to high-flyers, for example, would motivate them to achieve a greater degree of productive enterprise, and thus to produce more for everyone. The yardstick for this is a maximin calculation, looking to see whether the inequalities result in the best possible result for the least advantaged.

This calculation is made for several different areas, as follows:

Liberty. The claims of liberty have, for Rawls, an absolute priority after a certain level of economic well-being has been reached. This priority is a major keynote of the theory. For this reason it is treated separately below.

Above the level of priority, inequality is not allowed: trade-offs for increased economic well-being would not be accepted. Thus, the first principle, the right to the most extensive total system of equal basic liberties compatible with a similar system for all. This is reflected for example in Article 17 of the European Convention on Human Rights and Fundamental Freedoms 1950, which provides that the rights granted in the Convention cannot be used to deprive others of the rights in the convention. This has been the topic of case law, see *The Federal German Communist Party* case (1957) concerning the ban on that party whose objects were held to be incompatible with the freedom of association for others, and more recently the case of *Glimmerveen and Hagenbeck* v *The Netherlands* which concerned the outlawing of the expression of racist sentiments being held to be consistent with the freedom of speech.

It should be noted that Rawls is not here referring to the liberty of the individual but rather to certain liberties.

Fair equality of opportunity. Economic and social advantages are attached to offices and positions open to all in fair equality and opportunity. This is despite the fact that fair equality and opportunity might be contrary to the difference principle (see 'the difference principle' below, the maximin position on social and economic advantages). In some circumstances, one can envisage a 'closed shop' appointment system which might work better than an open system. Rawls argues that those who lost out in a system because fair equality of opportunity was not in operation would feel unjustly treated, and lose self-respect. (In fact that argument seems inconsistent

with some other things Rawls says: such feelings and considerations of desert wouldn't convince the self-interested parties. Also, the choice of fair equality of opportunity over available alternatives is not convincing, since criticisms against those alternatives, particularly that they are morally arbitrary, also apply.)

Just savings principle. We have seen the original position presumption that leads to this: the parties are taken to have inter-generational family ties. They are therefore concerned to save for future generations. This is the clearest example of a condition of the original position introduced, as a result of our considered judgments, in the reflective equilibrium process.

The difference principle. With the qualifications in the previous two paragraphs the parties will accept such economic and social inequalities as benefit the position of the least advantaged. In contrast with average utility, the least well-off do not sacrifice themselves for the better off; rather, they are put in the best position available. The difference principle, of course, follows logically from the adoption of a maximin strategy.

The main argument: why prefer the two principles to other possibilities?

It would here be necessary to outline briefly his reasons for rejecting various other conceptions, only some of which appear in his list. *Egoism*, as we have seen, is knocked out by formal constraints (requiring a choice of general principles).

Non-tolerant conceptions are knocked out by the veil of ignorance, because it hides the person's own conception of the good. By a non-tolerant conception, I mean one that would not be prepared to allow (certain) other conceptions to be pursued. For example, an extreme religious fanatic might have a conception of the good life that included a prohibition of other forms of worship: his conception of justice would include this provision. In the original position no one would choose that conception of justice, because there is a chance that they would be the holders of a prohibited religious viewpoint, an intolerable situation.

Non-tolerance will sometimes be linked with *perfectionism*, a conception under which society is organised in such a way that it furthers a particular ideal state of affairs. An extreme religious viewpoint might again provide an example. Such a conception would not be chosen, and this is another argument against non-tolerant conceptions being chosen, because the parties (not knowing their own conception of the good) have no reason to choose principles according with only one theory of the good. Instead they will choose one which allows freedom to differing conceptions in order that they, whatever their own 'good life' turns out to be, can pursue their own chosen course.

Rawls explicitly places *intuitionism* as a last resort. If any acceptable principles can be found by which society's affairs can be arranged, the intuitionist's suggestion that at each point the decision maker should follow his discretion can be rejected. Only if no such principles exist does the intuitionist win.

On the list of other conceptions, there are some *mixed conceptions.* These include the first principle with a different second principle. A particularly important alternative second principle is average utility plus a social minimum condition. Such a condition could resolve the problems average utility has in relation to sacrificing the least well-off. The two principles are to be preferred to this mixed conception, because of the difficulties of deciding the social minimum and the fact that it could be the two principles in disguise! The two principles would match its results, and should be chosen for their greater clarity and precision.

Apart from these various points aimed at showing why the parties would choose the two principles and not any other possibility, Rawls gives at this stage the considerations working in favour of the two principles. The first is that they do not involve sacrificing oneself for the sake of others, and therefore there are not the strains of commitment imposed by other theories. This is true of the least well-off, who are at their best possible position, unless of course they are envious (perhaps they will prefer to be poorer but equal?). High-level producers might not be so enthusiastic about a system which takes away their hard-earned profit to improve the position of the less well-off. Isn't the high-flyer being sacrificed for the bottom rung?

Second, the two principles have the advantage of psychological stability. Since no one is being sacrificed for others, and everyone's self-respect is enhanced, the two principles will receive public recognition and lead to a sense of justice. This 'advantage' is akin to the first; it is subject to the same sort of criticism.

The priority of liberty

The first principle is generally 'lexicographically prior' to the second; this means that its demands must be met in full before the considerations of the second principle are taken into account. (The exception is when the society has not yet reached a sufficiently developed economic position: this point is not explicitly decided by Rawls.) Liberty can only be sacrificed for liberty's sake in two circumstances.

1. Less extensive liberty must strengthen the whole system shared by all; an example is the rules of order in a debate.
2. Unequal liberty must be acceptable to those with less. For example, equality of opportunity might be sacrificed to lead to greater political liberty.

Rawls is not discussing all possible liberties, but a set list, political liberty, freedom of thought and conscience, freedom of the person and the right to hold personal property, and freedom from arbitrary arrest and seizure. The parties would not allow these freedoms to be limited for the sake of an improvement in economic or social well-being. Why not?

Rawls principally used the example of freedom of religious belief. Such belief is so central and important that the parties would realise that its sacrifice would be intolerable, and therefore is prohibited. This can be seen as a maximin point, that a poor economic situation and freedom to worship as you choose is always better than no freedom of worship.

A further argument utilises the fact that a certain level of well-being is required before the priority is absolute. Once that point is reached, survival and economic goods become less important, and cultural and intellectual pursuits (and therefore the freedom to pursue them) become increasingly so. Remember here the Aristotelian principle, that people desire to engage in more complex activities if possible.

Does Rawls establish the priority of liberty? First, he does not discuss sexual freedom (although at one point he says he has proved his point re 'religious and sexual freedom'); nor is it totally clear why the liberties on his list are chosen. Second, he only argues from religious freedom. It seems intuitively possible to distinguish between this and the other liberties. Religious liberty is central to the existence of most people (even if it is the liberty to have no religion) in a unique and most important way; the same is not true of the other liberties. Is owning a house as important as faith? And haven't people often given up their personal freedom for religion? The conclusion here might be that while priority is logical for parties to choose religious liberty, it is not necessarily so with the other liberties.

For these other liberties, the absolute priority of liberty over economic and social development is hard to accept, even given the fact that civilisation must have reached a level where economic needs are not desperate. Assume that by sacrificing five years' political liberty to a totalitarian regime, massive increases in wealth would result, including a great boost to the maximin position. Is it logical for the POP to rule out choosing this sacrifice? Wouldn't the resultant economic gain in fact greatly enhance the value of the restored liberty? In fact, Rawls does not find a proper place for the idea of the value of liberty: but what is the use of freedom of thought if the economic necessity to work proves too exhausting to allow it? And what use the freedom to attend plays/concerts if they cannot be afforded?

Finally, giving liberty this priority seems to me to be allowing into the theory Rawls' own conception of the good life. This point will be discussed more fully in the next section.

21.2 Evaluation of Rawls

There are many criticisms of Rawls' work. I shall consider some of the main ones that appear in the literature.

One criticism is that, through many avenues, his own conception of a proper plan for life, the proper way for someone to live, encroaches on the theory. The parties are denied knowledge of their own plan of life, their own conception of the good: as a result, society will inevitably be tolerant, even though the majority of people within it are in fact non-tolerant. The parties are motivated by a desire to maximise achievement on an Aristotelian scale of complexity, but it is not true for all. A monk, for example, would want no income, wealth, etc; many of the primary social goods would be useless for him. Rawls' original position fits only a certain category of person.

Further, an element is missed in the original position. The parties would logically be interested not only in how many primary social goods they got, but also how many everybody else got. This is not envy but a directly personal concern relating to quality of life. For example, imagine the parties deciding whether they should have cars. They would put into account not only the convenience of an individual having a car, but also the inconvenience and discomfort of living in a society where everyone had a car. This consideration is not mentioned by Rawls: surely it would affect the parties' determination to achieve maximum primary social good level?

Rawls is perceived as adjusting the conditions in the original position to ensure the principles that he wants are arrived at; behind a supposed veil of objectivity his subjectivity emerges. This is regarded as an intellectual sleight of hand trick, as he started from the two principles and their arrangement in lexicographical order and worked backwards. Simmonds takes a less cynical approach to this. He maintains that perhaps Rawls is 'attempting to elucidate the deep philosophical presuppositions that underlie his two principles'.

One of the most important criticisms of Rawls comes from one of his former students, Robert Nozick. Much of the following Chapter is devoted to an elucidation of Nozick's theory of justice, so it will suffice, I imagine, to demonstrate here some of his main criticisms of Rawls which will be further elaborated upon in the following Chapter. Nozick maintains that if one is interested in justice then one cannot deny the importance of Rawls' theory. One has either to work with Rawls' theory or explain why one is not working with it.

Nozick continues by attacking the notion of what he calls 'the patterned distribution of social good'. Any patterned distribution would really require us to consider the following problems:

1. If one is prepared to allow for the coercive redistribution of wealth then why not also allow the coercive redistribution of bodily organs? For example, if one has two healthy kidneys then why not give one healthy one to a person without any healthy kidneys, the point being that to regard ability or organs as common resources appears to give persons rights in other persons, so that the least advantaged have got a right to the best advantaged, raising the standard of living of the least advantaged. The example of taxation can be cited here. Where one person has to pay tax so that the money can be redistributed to the least advantaged then this is in effect a form of forced labour or slavery, because that proportion of one's time is spent working solely for others and not for oneself. It should be noted that Nozick is not talking about the need to raise tax in order to fund roads, hospitals or missiles, but about the tax that is used to pay for the poor such as supplementary benefits.
2. Nozick also rejects the difference principle as being unjust. He cites examples. Let us assume that everyone starts off with the same, an equal distribution. People will then freely enter into contracts and be prepared to pay in order to get what they want. In his Wilt Chamberlain example, to which we will return later,

Nozick shows that the spectators will become materially worse off while the player will become much better off. According to Rawls, that distribution is unjust because the least well off become poorer. Of course, this concentrates on the financial position of the parties and that may not be the entire picture. As we shall see, Nozick replaces this patterned distribution with his 'historical entitlement theory'.

Hart questions whether the parties would opt for the two principles and necessarily prefer liberty to equality of opportunity. Why not prefer equality to liberty? According to Hart, Rawls also underestimated the difficulty of balancing conflicting interests. It is the ideal that underlies his treatment of the allocation of resources. Rawls is interpreted by Hart as saying that a public-spirited person imbued with the notion of service to the community will decide never to give up any political freedom for material gain. Essentially this is because Rawls does not have any concept of the state in his theory. He pays too little attention to the institutional arrangements by means of which the distribution is to be carried out, in that those who in effect carry out the distribution will be the most powerful group in society. Can we really expect those people to act in a public-spirited way?

22

Robert Nozick

22.1 *Anarchy, State and Utopia*

Robert Nozick's *Anarchy, State and Utopia* is an important book, containing a new natural rights theory as well as criticisms of utilitarianism and Rawls, a plea for vegetarianism, and much more. It is recommended as a stimulating 'read' on justice and some related topics, though only if you have time and the inclination.

In brief, Nozick's theory is as follows. Man has certain natural rights, including the right to acquire property. These rights must not be violated by anyone, without the consent of the right-holder. They act as moral 'side-constraints' on action. To be justified, a state must be such that it would arise from a no-state position (the state of nature) without infringing the rights of anyone who did not consent; only a minimal state offering protection against violence, theft and breach of agreement would emerge in this way. Any further state is not justified; particularly, a state redistributing wealth is not justified, and taxation to bring this about is the equivalent of forced labour. The only legitimate way of coming to hold property is by just acquisition, just transfer, or rectification of a past injustice.

Nozick extols the virtues of eighteenth century individualism and nineteenth century *laissez-faire* capitalism. It has certainly represented a profound shock to legal theory. The book is a provocative essay and one which in my view has had a very considerable impact on political reality. After reading Nozick one may ascertain where many of the ideas of Thatcherism have derived their origin, although they have undergone some modification in the process. Nozick's views, to the extent that classification is at all legitimate, may be referred to as libertarian. He questions whether liberty and equality are compatible and concludes that they are not. His central thesis rests on the proposition that the individual is inviolable. This point is crucial to an understanding of his theory.

Let us look at his theory in more detail. In his critique of Rawls, as we have seen, Nozick rejects any 'patterned' conception of justice. A patterned conception is

one that views justice as a matter of the pattern of distribution of benefits and burdens that is achieved, for example, the Marxist idea of distribution according to need. As stated he prefers his 'historical entitlement theory', the content of which it would be appropriate to outline as follows:

Natural rights

Rather than examine the pattern of distribution, Nozick seeks to concentrate on the question of how the distribution came about in the first place. If that distribution is brought about entirely as a result of freely entered into transactions then it is just. He put it thus, 'If each person's holdings are just then the total set of holdings is just'.

The individual has certain natural rights, including the freedom from violence against his person, the freedom to hold property, and the freedom to enforce his other rights. Concentrating on the right to hold individual property, a person can legitimately acquire property in three ways:

By just initial acquisition

This details the circumstances under which a person may acquire ownership of formerly unowned resources. This right of appropriation follows Locke. Locke had the proviso that 'as much and as good be left for others'. Nozick has a more limited proviso, merely that the remainder be left for others, and not necessarily as much/as good. In any case, MacPherson says that the Lockean proviso ceases to be relevant once money is invented, since there is always some of that property available.

By legitimate transfer

This details the means by which ownership of resources may be transferred from one to another. If I choose to give you some of my property, or we agree to swap bits of our property, then you receive my property legitimately by transfer.

By rectification of past injustice

This details the action to be taken to rectify a distribution which is unjust in terms of the first two principles. If I acquire property in an unjust manner, it can be taken from me and restored to its proper owner. This principle in fact justifies a less limited state, in some circumstances, to remedy a series of past injustices. (Could it justify our present Welfare State?)

These rights cannot be violated without a person's consent; this is his meaning of the distinctness of persons. A person's separateness and individuality must be respected; he must not be treated as a means to an end. Each person has exclusive rights in himself and no rights in others. What is important is that in the pursuit of our own aims we do not violate the rights of others. As we have seen and stated above, Nozick's theory originated in a critical evaluation of Rawls. He has criticised Rawls on the grounds that individual abilities are not common assets to be exploited

for the benefit of the least advantaged. For this reason, Nozick rejects goal-based principles of justice. These are principles which judge a society by reference to whether or not it matches a particular goal, a particular end-state. Such principles will require the right of the individual to be sacrificed for the goal or desired end-state, the person being treated as a means to that end. The Wilt Chamberlain example, below, is a graphic illustration of Nozick's point.

Rather than such an end-state, goal-based principle, Nozick insists that a 'historical entitlement' principle be chosen. This means that a situation is judged not with reference to whether or not it matches a given end-state, but rather with reference to whether or not it came about justly, with no infringement of anyone's rights (hence the three just processes by which property may be acquired).

Under this entitlement principle, people's rights are respected: they become moral 'side-constraints' which forbid decisions and actions which violate them. Natural rights can only be infringed with the consent of the right-holder. For example, a road can only be built across someone's property if he consents to it: if he does not, his rights may not be infringed however much some particular goal (average utility, the best position for the least advantage, or whatever) may be enhanced by such an infringement.

These rights are the right to liberty and the right to property. Their inter-relationship is interesting. The right to liberty is defined by reference to the right to property and the right to property is the result of the exercise of rights in one's own labour. The right to property is then an expression of the right to liberty. Nozick believes that private property increases freedom, an idea that has considerably influenced Conservative politicians.

The idea of Nozick that, when one mixes one's labour with an object that is not owned, one acquires a right to that object which can then be transferred, does not address the question as to whether the exercise of one's labour gives a right to the whole value of the object with which it is mixed. What of natural talents and abilities? These are not possessed as a result of any labour but as a result of natural and, therefore, morally arbitrary distributions.

The minimal state

Nozick envisages a state of nature, and asks whether any state would emerge without harming people's rights. In fact, a state will emerge, through an 'invisible hand' process, that is, one which occurs without anyone intending it or aiming for it by morally permissible means and without anyone's rights (in Nozick's sense of the word) being violated. In brief, the process is as follows:

1. To protect themselves, people form protective agencies, pooling their protective resources and leaving themselves free from fear of attack.
2. In each region, one protective agency becomes dominant, but there are still independents.

3. The dominant agency will prohibit independents from enforcing their own rights, since they will distrust the independents' procedure for determining violations. This prohibition involves infringing the independents' rights (to enforce their other rights), and demanding compensation; this compensation is 'paid' by protecting the would-be independents as well.

The dominant agency develops into a 'night-watchman' state, carrying out a minimal range of duties, protection from theft, violence, fraud, and breach of contract. This state claims a monopoly of force. If the state engaged in a patterned distribution then it would be exercising excessive powers as it would entail constant interference with liberty. Nozick does, however, recognise the need for some state, otherwise there would be anarchy. Hence his minimal state can be seen as the way to Utopia where individuals are inviolate.

Distributive justice

Any state other than the minimal state is rejected by Nozick. He would clearly see the present UK set-up as unjust. Through social security and other aspects of the Welfare State, money is taken through taxation from the wealthier people, and given to the poor. This taxation is forced labour in disguise: such proportion of one's working time as is reflected in the national insurance and income tax contributions that are used for redistribution to the least advantaged is spent working for others. The redistribution involves violating rights to property. Unless it involves setting right past injustices (many would argue that it does) it does not fit into any of the three methods of just acquisition. Thus redistribution where resources are justly obtained would not rectify an inequality but would rather produce one.

Nozick rejects the difference principle. Social co-operation is a good thing, since it probably produces better results for everyone, but especially for the weak and poor who would have nothing if they had to act on their own. The difference principle, that economic and social disadvantages should be so arranged as to benefit the least well-off, gives all the benefit of their co-operation to the poor; which is asymmetrical and unjust. (This does not mean that Rawls' argument from the OP is wrong: it might be rational for the POP to choose it.)

The difference between Rawls and Nozick is in their starting points: Rawls starts from a standpoint of equality, and asks for reasons why we should accept inequality; Nozick starts from the idea of rights, with a consequence that a man owns the property he has worked for and created. For Rawls, the rich man must show why his wealth should not be taken; for Nozick, it cannot be taken without his consent. Wealth is created by individuals and they that create it have rights over it. Hence Nozick maintains that one is not entitled to regard society's total wealth as a cake to be divided up.

That this question of distributive justice is linked with that of goal-based versus entitlement principles is illustrated by Nozick's Wilt Chamberlain (a basketball star)

example. Assume, says Nozick, that at the start of a season your favourite end-state principle is satisfied in society: let us say, the difference principle. Chamberlain fixes a contract, giving him $1 of every spectator's entrance fee. At the end of the year, the million spectators who have watched him are each $1 worse off, and Wilt is a millionaire. Each $1 has been willingly given, a just transfer. Why should Wilt have to pay back some of his million to satisfy the difference principle again? Where is the justification for redistribution?

22.2 Libertarianism: Nozick's theory of rights

The opening argument of Nozick's *Anarchy, State and Utopia* consists of the following assertion:

> 'Individuals have rights, and there are things no person or group may do to them [without violating these rights] ... Our main conclusions about the state are that a minimal state limited only to the narrow functions of protection against force, theft, fraud, enforcement of contracts, and so on, is justified; that any more extensive state will violate persons rights ... and is unjustified.'

Consequently, the state may not coerce individuals to help others and may not protect an individual from himself. Such a view would seem to be diametrically opposed to that of the utilitarian. Rawls, who criticises utilitarianism because it does not give sufficient priority to rights, is criticised by Nozick for advocating that a priority in state considerations should be the minimisation of violations of rights. This is because, on the basis of Rawls' argument, an innocent man might still be punished in order that a rampage of vengeful citizens might be prevented from violating the rights of still more people. Nozick views the trading off of the rights of the individual in the interests of the rights of the many as a utilitarianism of rights.

Nozick's position is immovable. People cannot be treated as anything other than ends in themselves. The state and law must not violate the individual's rights, even if it is to avert the violation of the rights of others. The side-constraint of all state action is that the state's duty is primarily not to violate an individual's rights. Nozick's argument is that the individual cannot be forced into sacrificing his rights for the community, even if as a free agent he might sacrifice himself for another.

However, Nozick's consideration of utilitarianism is on the basis of the automatic assumption that it is not worth considering. Thus, his critique occasionally becomes absurd in the extreme. He says, for example, that maximising the average utility allows a person to kill everyone else if that would make him ecstatic, and so happier than average.

His attention to social, rather than individual concerns, is evasive:

> 'The question of whether [never violating the individuals rights is an] ... absolute, or whether [this principle] ... may be violated in order to avoid catastrophic moral horror, and if the latter, what the resulting structure might look like, is one I hope largely to avoid.'

Nozick's rejection of utilitarianism is thus on the basis that individuals have rights. What are these rights and how do they come about? The answer is that people (1) have a right to liberty, (2) have a right to the fruits of their labour. One person cannot have rights in another person. The function of the state is to protect the legitimate distribution of assets.

Liberty

Nozick cites Locke's idealist state of nature, where every individual acts as he sees fit, without leave or dependency upon the will of another man. Such individuals subscribe to the law of nature that no one ought to harm another in his life, health, liberty, or possessions. The function of the minimal state is thus to guarantee the laws of nature with minimal interference with individual liberty.

Nozick does not explain why people have the right to liberty. If a person is free to do as he chooses, then the question of rights has no real meaning. In a social situation, a claim of right is to assert that:

1. What the right-holder claims is his right;
2. It is to be preferred against the counter-claim of another.

Nozick is therefore asserting that the right of freedom is a right that prefers the individual's claim of freedom of action, against another's claim that he might restrict that freedom. Yet is this not to contradict the freedom of another to act? If my freedom to act in the way I choose is to be preferred to your freedom to choose to restrict my action, do I not therefore assert that there are limitations on your freedom of action? This would contradict the assertion that all people have the right to act freely. This Nozick explicitly accepts by postulating that natural law states that we cannot use our freedom of action to endanger the life, health or possessions of another. The law is justified to compel obedience to these conditions.

This would seem to solve the problem. However, if I have access to food and you do not, yet I refuse it to you, do I not do as much harm as I would have done had I stolen your food from you? Freedom in a social context depends on the absence of monopoly. We have seen that Nozick criticises the utilitarian on the grounds that he would allow one man to kill everyone for the maximisation of his pleasure. Yet Nozick would allow by default one man, in the name of his individual liberty, to allow all others to starve!

The individual in society, or even in competition, is dependent on either the compromise or weakness of his fellows for his own survival. Since Nozick rules out physical violence as in violation of the rights of others, man is thrown back on the distributive justice of his fellow men.

A further point to note is that a man may become comparatively freer in his opportunities, if he sacrifices his complete autonomy and co-operates with another. In modern society, the benefit of communal action is manifest. One man cannot in his lifetime build all the things that he consumes or uses in his everyday life.

However, neither can I enter into a voluntary transaction with all those from whom I might benefit. Nozick seems to deny non-consensual, co-operative action. If the value of a man's contribution to a particular goal is greater than the benefit that he, as an individual, gets from the common pursuit of that goal, it would be wrong to force him to engage in it. Yet people are unable to gauge the benefits that they get from society in any direct way, and are not always the judge of their own best interests. Nozick simply views people's interests at their own concern. However, in a society people's interests are interdependent. My well-being is not only my own concern, but my family's, my employer's, and those people I am philanthropic towards. I am an individual only because I have a discrete and valuable place in society.

Property

The cry of anarchists such as Proudhon and Kropotkin is that 'property is theft!'. The basis of this is that, by asserting that I may have a sole right to an object or commodity, I may be denying another what he needs. The true libertarian must presuppose that no man has the right to deny another his liberty. A fundamental precondition to liberty is that a person must be alive. Thus, a precondition of liberty is the satisfaction of the basic needs of life. Basic commodities and means of survival must be distributed not according to the means of acquisition, as Nozick suggests, but on the basis of the needs. Most libertarians therefore assert that the right to liberty determines that the individual has the right to the necessities of life, by which he may obtain liberty. Nozick believes that the free market naturally will tend to the satisfaction of needs. He presupposes that minimum needs are satisfied, so that law should not interfere to ensure it.

Conclusion

Lukes says that Nozick fails to appreciate the nature of the individual as a social being. In terms of supplying us an answer to the problem of the nature of the legislators' duty, his theory simply endorses the rights of those who have already got rights. He endorses the inequalities in society on the weak assertion that to interfere would be to damage the rights of some in order to benefit those that have no rights.

22.3 Evaluation of Nozick's theory

Nozick's theory is interesting and a strong challenge to Rawls. It is not, however, without its defects. What becomes clear from an evaluation of this topic is the close relationship that the authors bear to current political agenda. A major evaluative point in favour of Nozick is that many of his ideas are now the topic of intense political debate both in this country and in the United States. In that light perhaps we ought now to examine some of the main criticisms that have been made of his theory.

He takes rights as his starting point. The rights he takes are of uncertain pedigree: how are they derived, where do they come from? There is a strong body of opinion that denies the possibility of any objective rights, such as those Nozick must contemplate. He sees his rights as inherent, as natural. The idea of such rights should be established (if it can be), not assumed. Further, the choice of rights to assume is a value-laden process; why not include a right to welfare and help from others? Or, to put it another way, why limit the 'moral landscape' to just rights? Why not have duties as well, to assist others in need?

The argument from rights to a minimal state is not without its problems either. Particularly, at one stage, the dominant protective association prohibits the use by independents of their own rights-enforcement procedures, at least unless those procedures have been vetted and found acceptable. Surely this involves a strong violation of the independent's rights to protect himself, which is not properly compensated for even by membership of the dominant agency? And wouldn't this compensation be free, resulting in only some people paying for the services?

In any case, it is not clear that the dominant agency idea would work out as Nozick thinks. Might not the strong people in a community think that the only opposition or problem will be from other strong people, and therefore form an organisation with them? It would be cheaper for all concerned if weak people were not allowed to join, and prohibiting their own enforcement would not be necessary, because it is not to be feared (since they are weak). Doesn't the state idea therefore include some element of compassion for the weak?

On the other hand, assuming that we get to a minimal state, does not practical sense require even the right and strong to accept a greater state? Assume you are rich and strong now, and tomorrow you are robbed, or lose all your money on a business deal, or are knocked down and paralysed. And don't most people grow old and infirm? Isn't it practically minded to accept and agree to pay for a state which will provide a safety net in case such things happen to you, or your family and friends?

Lukes argues that Nozick has an unreal conception of the relationship between the individual and the society, to the extent that he has excluded what for Lukes is the ever growing role of the state. Lukes further maintains that the central flaw in Nozick's arguments is 'the abstractness of the individualism they presuppose'. Lukes maintains that it is not possible to divorce an individual from his society.

Hart, not surprisingly, takes issue with Nozick's theory. He maintains that Nozick's assault on utilitarianism is paradoxical as it shows that he is unwilling to disturb the exisiting pattern of distribution. Hart further takes exception to Nozick's likening of taxation for the sake of redistribution of wealth to slavery, which is so rooted in Nozick's belief in the absolutely inviolable character of property rights. Hart maintains that one is talking of two different types of burden. Man is free to decide whether to work, what work and how much work to do; a slave is not so free. Hart would speculate whether rights which are derived from human interests and needs could outweigh property rights.

Lloyd and Freeman are unsure as to how Nozick's minimal state would emerge from the state of nature without infringing individual rights. They say that Nozick does not adequately explain this. Furthermore they maintain that Nozick leaves many questions unanswered and makes a number of assumptions which do not stand up to examination, not least the point of where do people get their rights from? Is this historical, some initial act of appropriation which then confers unlimited rights of use and disposition?

Nozick's rejection of welfare rights neglects the interests of the weak. Nozick's view of the assistance to the weak is privatisation of philanthropy. The matter is essentially one of private charity, a view echoed by one British government minister following the recent cutting of the upper tax levels, when it was suggested that some of the money handed back to those taxpayers ought to go to private charity. Human dignity and the receipt of charity are not clearly compatible.

23

Ronald Dworkin's Theory of Justice

23.1 Liberalism's foundations

In Chapter 7 we have considered in detail Dworkin's scheme of equality of resources. We have seen that it is guided by the important foundational – humanistic – principle that people should be treated as equals. But we have also seen that there is a need for a fully worked out division between the public and private responsibilities of people. We have seen that distinction worked out by Dworkin in a number of different contexts. It is inherent in his rejection, in the ideal world, of utilitarianism, even of the egalitarian sort. Dworkin develops the idea of the division in an attack on the contractarian account of liberalism in his Tanner lectures of 1990. But it is necessary, first, to understand what is the substance of recent debates on the justifications for liberalism, and second, to understand the nature and significance of the contractarian approach.

Dworkin's own liberal theory is best understood by considering his distinction between what he calls 'discontinuity' and 'continuity' theories of the foundations of liberalism. The contractarian theory just described is a discontinuity theory. The discontinuity is between a person's personal ethics – what Dworkin describes as 'first person' ethics, or 'well-being' – and 'third person' ethics, or 'morality'. Thus, contractarian theories allow for first person ethics to provide the justification for the existence of the contract, and for third person ethics to be justified only by reference to the contract itself.

Dworkin says that the paradigm for a social contract theory is an ordinary commercial contract. There are different personal reasons why we might enter into such a contract but the rights and duties are established not by those reasons but by the contract itself. It means that the contract acts as an artificial social construct from which rights and duties flow.

Dworkin cites Rawls as having the most sophisticated version of the continuity strategy. Rawls' view is that the basis of liberalism must be sought in an 'overlapping consensus' among different comprehensive ethical views. In other words, at the basis of liberal political principles could only be a shared assumption that these were required in order to provide for co-operation in society where there were different ethical views. Such liberal principles are not, as Rawls has said, to be thought of as a mere *modus vivendi*, that is, as necessary to ensure self-interest, but as a moral basis for liberalism. People should come to see that the liberal principles connect to each person's different moral interest.

But the rights and obligations that people have, under Rawls' scheme, derive from a perspective that is not personal, because it is founded in the idea that people of different convictions about personal ethics should endorse liberal perspectives for reasons other than those to be found in their personal ethics. That, at least, Rawls is clear about: that the political principles of liberalism are not to be drawn from any comprehensive theory.

23.2 Liberalism and personal ethics

Dworkin's project, on the other hand, is to make a bridge between personal and political ethics, so that ethics is part of liberalism's foundations. He agrees that the personal perspective is everything the liberal political perspective is not. We are not neutral and impartial, as liberalism claims the state should be, but committed and attached.

The distinction between the continuity and discontinuity strategies is helpful. It serves to distinguish the theories I mentioned earlier. What is the status of the confused argument for liberalism from the basis of the supposed subjectivity of moral judgments? It purported to derive neutral principles from the nature of moral argument itself. Put at its best, buried in its premises is a proposition about mutual respect: since I can no more prove my moral assertions to be true than you can, our moral assertions shall have equal weight. If this is so, then it is a continuity theory.

But another interpretation might find tolerance to rest upon self-interest. We could say that because of the non-provability of moral propositions, we should deal with others at arm's length to ensure that they keep at arm's length from us. That interpretation makes out the connection between one's personal ethics and the political perspective to be discontinuous.

What about the hippy liberalism to which I referred? That clearly is a continuous theory because it requires you to adopt as part of your personal ethics the personal ethics of everyone. That is why it is confused. Nevertheless, its strategy is to encourage everyone to extreme tolerance through the development of each person's personal ethics.

Raz's theory is a continuity theory, too. The political perspective is defined by a personal ethics which places very great weight upon the principle of personal

autonomy. That principle must be endorsed by the state in the form of tolerance of a plurality of different exercises of personal autonomy. In other words, state tolerance follows from a personal ethics placing great importance on personal autonomy.

The appeal of liberalism

Let us now examine Dworkin's continuity version of liberalism which, for the political sphere, he calls 'political equality'. Dworkin says that there are three major problems which any theory of liberalism will have to face, those of the visionary appeal such theories should have, the promise they have of attracting a consensus about them and, in particular, what he calls their 'categorical force'.

These problems arise from the difficult problem in liberalism of drawing a line between the personal and political perspectives. In the first place, since it is *premised* on the idea of people having different views in their personal ethics, on what constitutes the good life and personal well-being, how can any such theory hope to have either visionary appeal or consensual promise?

At first sight, the premise of difference seems to deny the possibility of people being equally struck by an equality vision of the future. What appeals to one person's personal convictions will not appeal to another. Further, if that is the case, what is the hope for being able to attract a consensus about the idea? Dworkin goes so far as to say that it is unrealistic to suppose that political liberalism can gain a consensus yet.

There is a particular problem in the idea of the categorical force of political liberalism. From what basis can we claim justification for the moral strength of the principles? If people have different ethical views, views which may be partial and committed to different forms of life, how can independent moral force be accorded to the neutral and impartial liberal principles?

These three questions, concerning visionary appeal, consensus and categorical force, usefully throw into relief the difference between the continuity and discontinuity strategies. Take, for example, the visionary appeal of a discontinuity theory. Since it assumes different ethical views, it is difficult to see what visionary force it can have. Its appeal is not supposed to lie in any person's particular personal ethical perspective. In what, then? The best will be in some idea, such as Rawls', of 'mutual respect and co-operation' but, as I have already pointed out, that idea is only hopefully something beyond self-interest (and self-interest, we assume, is not the same as personal ethics).

A similar problem arises with categorical force. From what moral perspective does this come? The short of it is that no one is going to accept as binding upon him a proposition which is not part of his personal ethical perspective. The discontinuity strategy assumes different ethical perspectives, so that the categorical force can only arise from the contract. Here we can return to the paradigm of the ordinary commercial contract, in which the rights and duties flow from the contract and not the personal perspectives of the parties.

Dworkin claims that it is an 'insane' theory that there was ever actually a contract between citizens to form a state. No obligations of the commercial contract sort arise in advance of the contract coming into existence, not even where such a contract is clearly about to be made. The categorical force, then, of discontinuity theories cannot be located within the structure from which the rights and duties are supposed to flow. The discontinuity theory fails, because it cannot account adequately for the three features which he, reasonably, says are necessary for a successful liberal political theory. But what about the continuity theory? This fares much better on visionary appeal and categorical force. Both are derived from the personal ethical perspective, so that a person can be called upon to endorse from his own perspective the political structure proposed. There is no need to pray aid from an intermediate stage similar to the commercial contract.

Of course, many will have difficulty with this idea in Dworkin, because of the supposed problem of objectivity. What hope is there, ever, that the categorical force for liberal principles can arise from personal ethical perspectives? Perhaps moral views are too radically different for the continuity project to be successful.

There are two ways in which we might take this objection. First, it could merely express the simple view that moral argument, say, about the categorical force of political principles of liberalism, will always be controversial. But, for Dworkin, lack of demonstration of truth is not an obstacle to moral argument (see Chapter 25, on the 'one right answer' thesis).

But second, it could express a more complex view that the controversy is such that *argument* is not possible, that the incompatibility is of a nature that not even modes of argument are shared. This view may be that of people who talk of the presence of irreconcilable conflicts (a 'contradiction' of principles, as opposed to 'competition') in society. They mean that conflicts exist which are simply not resolvable. This view is more sophisticated than the 'subjectivist' view, because it can allow for moral controversy, but asserts that, in some societies, the controversy can be irresolvably deep, where arguments necessarily pass one another.

So, the situation remains, for Dworkin, one where there *is* the possibility of argument, and the categorical force of the political liberal principles will spring directly from the (right) personal ethical perspective. It is not an argument against *this* view that people might disagree, therefore. To point to that fact *that people will disagree* is merely, in his view, to make a statement of the sociology of moral argument.

The categorical force of liberalism

But how can categorical force derive from the personal ethical perspective really? After all, our personal perspective is coloured, as Dworkin points out, by ideas of partiality and attachment, not the ideas of impartiality and detachment required in political liberal principles. It was, in fact, this apparent contradiction which, in his view, sent the contractarians to formulate their discontinuity strategy.

It is instructive to look at two ways of answering this question, each of which Dworkin believes to be unsatisfactory. One is that we could, using the idea common in political philosophy, say that in matters of state, the 'right' takes priority over the 'good'. In other words, perhaps in politics where, as Dworkin says, 'the stakes are higher', morality should be given exclusive force over personal well-being.

But he does not think this argument is sufficient. It is no more than an assertion because it offers no independent reason why morality should have sovereign force in the political sphere. Particularly, he says, the ideas of personal well-being and morality are not independent. The idea of fairness, for example, allows partiality towards friends and family, so that within the idea of fairness, as it were, is the idea that friends and family are important.

So there may be a bridge between the sorts of principles he wishes to endorse at the political level, and the personal ethical perspective. The connection requires a deeper analysis of the relationship between fairness, justice and impartiality, on the one hand, and the personal perspective, which Dworkin later gives. But it is not enough, he thinks, simply to insist that, in the political sphere, the right takes priority.

23.3 Our personal ethics

Dworkin claims that a person should be free to use his resources as he wishes. In other words, a genuine understanding of treating people as equals (the 'principle of abstraction' operating through the 'bridge principle') means that invasions of liberty are invasions of equality as well. Of course, invasions of liberty will be justified where that is necessary to protect an egalitarian distribution of resources and will include, for example, the protection of personal security. Invasion of personal liberty on other grounds, such as intervention to influence private sexual behaviour, will not be justified, however.

The political equality implied by equality of resources means that only equality of 'outcome' is justified, not equality of 'impact'. It follows, says Dworkin, that democracy is defined by outcome as well as by other things flowing from the worthwhile life. He sees political activity as flowing naturally from personal moral experience and this idea is his answer to the charge that his theory lacks the dimension of community. Democracy means more than just the formal opportunity to vote. It requires the much richer idea of politics as a theatre of moral commitment and debate.

The idea of our 'critical' well-being

Dworkin claims that what is important about one's life is constituted not by what one *wants*, what he calls 'volitional well-being', but by one's 'critical well-being'. The difference is this. 'Critical well-being' is what you should want as opposed to what you actually want, which is 'volitional well-being'. The idea of the satisfaction of

wants, such as pleasure, is too unstructured and insufficiently complex to explain the judgments we make about what is good in life.

Dworkin's analysis here is obvious. We do not think that what is worthwhile is *constituted* by simple want satisfactions of whatever kind. As he says, your life is not better because certain of your wants, such as being able to sail better, are satisfied, or your life is not worse because you suffered in the dentist's chair. There are some wants, however, which do matter for your life in the relevant way. Dworkin suggests one such as *wanting* to have a better relationship with your family. And the only way to distinguish between those wants that are important and those that are not is by abandoning the simple idea of want satisfaction, or 'volitional well-being', as constituting what is important.

Dworkin does not think that this distinction commits him to a distinction between 'subjective' and 'objective' wants. Clearly, it does not, even granted the general unease his methodology has about the meaning of that distinction. But, to employ the terminology, 'subjectively' I can distinguish between those wants I have that express volitional interests and those that express critical interests. Dworkin's distinction is only a graphic way of distinguishing between what I consider important and what I do not.

The good life constituted by performance

We can be agreed, then, that what is important to a person's life is what is important as judged from the perspective of that person. Dworkin goes on to draw some very useful distinctions. First, he distinguishes between the 'product' value of a life, measured by what that life produces, and a life's 'performance' value, measured by how a life is lived. A life of good product value would be something like Mozart's life, because he produced great works of music, or Alexander Fleming's life, because he discovered penicillin.

A life of good performance value, on the other hand, would be one where a person responds to his circumstances in, as Dworkin says, an 'appropriate' way. We can see what he means. A composer might live a life both of performance and product value. We might say that his life, lived as a performance, achieves value from the *way he lives it*, quite apart from a judgment that *what he produced* is of value.

Employing the idea of critical well-being, Dworkin now says that idea makes most sense only on a judgment about the performance value of a life. If we only judged lives according to their product value, as he says, most of our lives would be 'puny' compared, say, to that of Mozart or Fleming. Under the performance model, however, the response to life is parametered by each person's particular capacities. The goodness of a life is not judged in the shade of that life's 'product' but in terms of how it has been lived. Under this account of critical well-being, the brilliant person produces something better than I do, but the value in his life is measured against his response to his circumstances. You can see that here Dworkin is drawing upon the same sort of argument that led him to devise a tax on talents. At root, it

strikes at rewarding people for what is merely a matter of luck. It is true, of course, to the liberal tradition of regarding this sort of *desert* as having no place in the distribution of resources.

The parameters of the good life

What is the role of endorsement of what is of value in your life? Dworkin suggests that the important idea for critical well-being is that the good life is one that you endorse. Further, it does not make terribly good sense to say that your life has value without endorsement, because the good life is one that is constituted by your doing what you critically believe you ought to be doing. In other words, endorsement is 'constitutive' of leading a good life.

What would be an alternative? Dworkin suggests that endorsement could have value in its being 'additive' to whatever else is of value in a critical life. But, as you can see from my previous argument, that idea is going to have very little value on the performance model if, indeed, it has any meaning at all. Dworkin says the additive view fits the product model much better. A person who values what he has produced has additional value in his life, thereby, than someone who does not.

The idea appears bizarre for the product model. Dworkin drops it, by reason of his preference for the performance model. Let us say that Mozart's life is better for his having endorsed his products, as it were, than if he had merely produced the works for money and not thought of them as otherwise worth doing. Does that make much sense? Does it really make any difference to the value we place on Mozart's life whether he valued what he produced?

Perhaps it is another way of saying that the 'added' value could only come from judging a life by its own internal convictions, which is to confirm Dworkin's analysis of the performance model. He gives the example of Alexander Fleming's janitor who disobeys his instructions and omits to throw away the mouldy culture dish from which Fleming later discovered penicillin. Fleming's janitor's life had product value. Would it make the slightest bit of difference that he had endorsed his breach of duty?

23.4 Justice and personal ethics

Now we come to the conclusion of Dworkin's lengthy discussion of philosophical ethics. Do we say that a person's life is critically good measured against the resources he actually has? No. We can say that a person's life was not a good one *just because he had too few resources*. Here is the nub of the argument. It is now clear where Dworkin is going:

> '... the best life for a particular person, we might say, is the best life he can lead with the resources that *ought* to be at his disposal according to the best theory of distributive justice'.

This is an arresting conclusion. It means that justice enters ethics by limiting the amount of resources a person can have to live a good life because, you will remember, the measure of equality of resources is its true cost to other people. Dworkin claims here that there is in this idea a shadow of Plato's claim that justice is always in a person's interests.

What is congenial about Dworkin's conclusion is that the answer to the question of whether I have lived a good life ties the question of the ideal performance of my life to that of justice. For example, I cannot regret not having done well at politics if it is clear that I could only have done so had I had an unjust amount of resources. Or, says Dworkin, I cannot be (rightly) pleased by having lived well despite my having had only a pauper's share of resources.

23.5 Evaluation

We are now in a position to see the connection between personal ethics and the claims of liberal equality. Justice enters the personal sphere because it sets the parameters of that sphere. Our own ethical life is coloured by the justice of the distribution of freedoms, particularly in the area of resources. Crudely, I cannot escape the effect justice has on my own ethical life, measured as a matter of critical performance and from my own point of view. Justice and my personal ethics cannot be separated. That means that a proper concern for my own personal ethical life must lead me to a proper concern for the just distribution of freedoms in the community. The continuous link between personal ethics and the political structure is, in Dworkin's view, thereby established.

It may be instructive to see how far this conclusion departs from the contractarian line which Dworkin sees himself attacking. Let us go back to what he sees as the paradigm of the contractarian argument, which is the ordinary commercial contract. The contractarian line says that the rights and duties arising under the contract are independent of the personal ethics of each party. That is how the contractarian justifies the neutrality of political liberalism without any inconsistency with personal ethics.

But, if we are to employ the metaphor of a contract, and it is no more than a metaphor, then we can see that it only makes sense to talk of contracts if we endorse the institution of contract-making and contract-enforcement. We see the *ethical* sense of them. (We *need* not. We *can* claim that contracts are merely institutions which in the long run fulfil our non-ethical self-interest.) In other words, it is not too difficult to give personal ethical justification to the contract's neutral way of distributing rights and duties.

These remarks are not critical of Dworkin's project. His attack is against a line of specific contractarian thinking that really does regard the line between personal ethics and political liberalism as genuinely discontinuous. Despite his generous efforts to make the best sense of Rawls' idea of an 'overlapping consensus', in looking for an interpretive account of justice 'deeply embedded' in a pluralistic community,

Dworkin does not succeed in rescuing him. Simply, the way Dworkin denies the contractarian line is to say that the contract can only make sense seen as a striking way of showing that the sense of justice which sets the limits to our personal ethical convictions is the same sense which orders the principles of liberal equality. Dworkin's theory appears to be the only one at present touching these deep issues.

Judicial Reasoning

24

The Common Law Tradition

24.1 Introduction to precedent

The judge has two tasks. He must resolve the dispute before him and he must reach his decision by reference to some impartial rule of law. One of the most obvious aspects of formal justice is that all cases should be treated alike; one of the commonest and most noteworthy features of many human institutions (clubs, societies, companies, etc, as well as states) is the tendency to repeat earlier practice and follow earlier patterns. For these and other reasons, most legal systems have developed a system of precedent, including the use of past decisions as a guide to present decision. A moment's thought by any student of English law should bring scores of decisions based on precedents to mind.

As we shall see, though, the English system – the common law system – in fact uses precedent in a slightly different way from civil law systems. In England, precedents of an appropriate authority not only guide decisions in later cases, but bind the judges in those later cases: within the given hierarchical structure, a judge in an inferior court may obey the decision of a higher court on the same point. This is the doctrine of *stare decisis*.

Any system using precedents will require a method of keeping them in an acceptable and accessible form; this need for law reports is obviously greater where precedents are law (since they bind later decisions, they are actual law, and not just guides to what the law is). In the *stare decisis* based system there will also need to be a defined hierarchy, and an established way of working out what part of an earlier case is binding: we call this the *ratio decidendi* (reason for deciding).

In the following sections, we look at the doctrine of *stare decisis* and how, if at all, it differs from the civil law use of precedent; the flexibility introduced into *stare decisis* in various ways (including an analysis of the problems involved in identifying the meaning of *ratio decidendi*); and the present English rules on precedent and *stare decisis*.

24.2 *Stare decisis* in theory

From the Latin *stare decisis et non quieta marere*: the doctrine of *stare decisis* lays down that decisions of superior courts bind the lower courts in later cases. The exact details will be discussed in paragraph 24.4. The courts fall into a hierarchy, House of Lords (HL), Court of Appeal (CA), Divisional Court (DC), judges of the High Court (HC) and so on, with the European Court of Justice (ECJ) thrown in for good measure.

This doctrine is the result of a combination of historical factors, really beyond our scope. One necessary factor, as was noted in the introduction, is a satisfactory system of law reporting. From the Year Books onwards, law reporting in England has been a developing and now integral part of our court system. The private collections of law reports (for example, Cokes) gave way in the nineteenth century to the reports of the Incorporated Council of Law Reporting whose reports remain the most authoritative (since they are checked by the judges), although still unofficial. There are many other series: reports in *The Times* each day, the weekly *All England Law Reports* and the *Weekly Law Reports* (published by the council in addition to their main series), and many others. The computer revolution has made possible the storage of details of many more cases; the HL has recently disapproved some of the consequences in terms of increasing citations of cases from computer records.

Past cases – once the *ratio* is determined above – bind. This distinguishes the English doctrine of *stare decisis* from the treatment of precedent in civil law countries, and even in some common law countries. In civil law countries, past decisions are not binding, but merely persuasive, with the strength of persuasion depending on the authority of court and judge; in some common law countries, for example, the United States, the *stare decisis* doctrine is not applied as rigidly as in England. Several factors play a part in this distinction. In both the United States and France, for instance, the court structure is not as strictly hierarchical as in England, with many, particularly state and district courts, of concurrent jurisdiction with no authority over one another. Further, the basis of French and other civil law

is a code: past decisions must always be justified on the basis of the code, and its wording can always provide a justification for not following past cases. Similarly, with US Supreme Court decisions on the Constitution; the Supreme Court's role as the arbiter of that document mitigates against strict *stare decisis*, and many of its landmark decisions (for example, *Brown* v *Board of Education* (1954) outlawing segregation in schools) are in fact reversals of earlier rulings.

The mode of reporting and of giving judgments in France also works against the English model. Judgments tend to be pithy statements, frequently only on the facts; they are often accompanied in the reports by influential commentary on the case and its effects by jurists. This tends to decrease the role of the judgments, and increase the importance of the learned writings, in discovering what the law is.

In any case, the differences between France, for example, and England in this respect can be over-estimated. While the code is the last word, it is – just as English statutes are – often uncertain or vague, and it is the decisions of the courts which make the detailed law. A set of decisions pointing the way will, in France, be quite settled; in England, one decision on a particular point may be binding on lower courts, but certainly not on upper courts. Also, the authority of the French Cour de Cassation is such that its decisions are almost always final.

In marginal cases, there is a difference: even a long series of cases does not fully bind a French judge, and particular precedents considered incorrect or out of date or unjust can be overruled or not followed without fear of criticism by higher courts, and without some of the devices (such as distinguishing) we discuss in relation to our own system in paragraph 24.3 below.

Does the doctrine of *stare decisis* have a value over and above the ordinary precedent system? The advantages held by the latter are certainty (to enable people's affairs to be arranged and conducted within a known legal framework), uniformity (like cases treated alike) and logic (fields of law developing harmoniously), mixed with a degree of flexibility to prevent injustice. The ordinary system sacrifices a degree of the certainty, uniformity and logic of the *stare decisis* system for the benefit of slightly increased flexibility and, hopefully, decreased injustice.

Think about the two options, bearing in mind the various tones of flexibility introduced into the *stare decisis* system. Is either option clearly the better one?

24.3 *Stare decisis* in practice: flexibility

The bald statement of the *stare decisis* doctrine makes it appear rigid and inflexible. In fact, in practice, judges do have a wide measure of flexibility and movement. If a judge does not want to follow a particular precedent, there are several techniques or devices he can use to avoid it: such avoidance is not always possible, but it frequently is.

One factor a judge always has to weigh up is the authority of the report itself and of the court. Present sets of reports are generally considered accurate (although the

councils reports are the most acceptable, as they have been checked by the judge), but earlier private sets of reports were not complete, and are of varying quality: Coke's, for example, are thought to be of high quality. Since the hierarchy is so important, a judge must always decide if he is bound by the cited decision or if it is just persuasive. If it is just persuasive (Privy Council, lower courts, other judges of the High Court perhaps, foreign judgments), the judge must weigh how much persuasive authority it has (Privy Council judgments, for instance, since they are normally given by House of Lords members, are very persuasive).

A judge must then decide which parts of the earlier case actually bind him. He must distinguish the *ratio decidendi* of the earlier case from the *obiter dicta* in it (which do not bind him); and this distinction is one of the major sources of flexibility. While any student will quite happily expound on the *ratio* of a past case, and be prepared to inform a judge of exactly the extent to which he is bound, in fact the actual definition of a *ratio decidendi* is uncertain: and frequently it is difficult for a judge to identify the correct *ratio*. Is there any definition of ratio decidendi that adequately captures judicial practice?

Definition of ratio decidendi

The traditional view of *ratio* is that it is the rule of law enunciated by the judge to the extent that it is necessary for the decision of the case. Even if we do think that the judge's expressions of relevant law are the *ratio*, this definition is not practically very useful: the important question is, what part of the judgment is relevant? In *Donoghue* v *Stevenson* (1932) was Lord Atkin's neighbour principle relevant and necessary, or just the narrower principle relating to manufacturers liability? Also, what if the judge does not state the law, but just decides the case before him? In any case, the statements of the judge are not always considered to be correct statements of the *ratio* when considered in later cases: it is not thought doctrinally incorrect to say the case really decided X, even if the judge said Y.

If we reject the traditional view, we find no shortage of suggested alternatives to take its place. Wallbaugh proposes a reversal test: if the reverse of the proposition would have led to a different decision in the case, that is the *ratio*. However, that does not help us distinguish between the two statements of principle in *Donoghue* (since it is not clear for which of them the Wallbaugh test is true), nor between them and general statements such as, there is a tort of negligence, manufacturers can be liable for negligence, and so on. The reversal test can tell us what is *not* the *ratio*, but cannot help us work out what *is*.

Others will argue that one should try to find the underlying principles. At what level of generality? Also, it is acceptable to reject the underlying principle of a case like *Donoghue* (the neighbour principle, perhaps?) while considering the case to be correct on a narrower ground (the manufacturers' liability). The underlying principles test is too vague. Two other tests, which we can quickly reject as being contrary to our experience of how judges work, are those of Lord Halsbury in *Quinn*

v *Leatham* (1901), that a case is only authority for the order made on those facts (this seems far too narrow to capture the width given to the *rationes* of past cases), and of Lord Devlin, that the *ratio* is the reason for the decision which the judge wishes to be the source of precedent. (Is it then incorrect to say that a case is a precedent and binding in a way the judge never intended?)

A definition which has carried much persuasive weight is that of Professor Goodhart, for whom the *ratio* is the decision based on the facts treated as material by the judge (he was particularly concerned to move away from treatment of the *ratio* as the judge's statements of law). A judge views certain facts, explicitly and implicitly, as material: his decision on those facts is the binding *ratio*. This view is interesting, and expounded at length by Goodhart (see Lloyd and Freeman's *Introduction to Jurisprudence);* but some problems do arise. It is often difficult to tell which facts the judge implicitly takes into account, and *ex post facto* any interpretations thereof may well be wrong; while there is always the problem of being tied to the facts the judge found as material.

The approach of Professor Stone is illuminating. He maintains that there is not a unique *ratio* of a case, but rather a choice of *rationes* available for later judges to choose from. Stone identifies two possible *rationes*, the descriptive and the prescriptive. This, it is submitted, is a good explanation of the nature of the common law system. The descriptive *ratio* is ascertainable from the decision once given, but the prescriptive *ratio* is how a subsequent court treats the earlier decision. In *Evans* v *Triplex Safety Glass Co Ltd* (1936) where a windscreen smashed and caused injury to the driver of the vehicle, the court – bound by *Donoghue* – held that the *ratio* of *Donoghue* was that a duty of care arose only when there was no possibility of interference with the product between the time it left the manufacturer and the time the loss was caused. The court held that there was such a possibility in *Evans* and so the plaintiff would not recover. The view of *Donoghue* stated in *Evans* was the prescriptive *ratio* of *Donoghue*. Dias goes slightly further, and suggests that the *ratio* should be viewed in a continuing time framework, as the interpretation of the case given by later judges. These views help us to understand a central feature of the *stare decisis* precedent system, that it is important to see how cases are treated in later cases to discover for what they are taken as authority: in *Donoghue*, the example we have been citing, it is clear that it is authority in 1988 for the neighbour principle.

However, the Stone, or Dias, view does not provide us with a definition which explains how the judge decides what the *ratio* of a previous case is: in the case of negligence immediately following *Donoghue* a judge had to decide what its *ratio* was. Knowing that there were several for him to pick from (Stone) and that the full import of the case would not be known until after the series of decisions (Dias) doesn't make it easier for us to understand the use of *Donoghue* made by that next case judge.

Montrose has stated that the argument is essentially one of a terminological nature. His purpose was to reassert the strength of the common law tradition. He seeks the meaning of the *ratio* and identifies three possibilities:

1. the rule of law to be found in the actual opinion of the judge forming the basis of his decision – this is the meaning that Montrose preferred;
2. the rule of law for which the case is binding authority;
3. any reason which ultimately brings about the decision – essentially this relates to the reasons for the *ratio*.

Does this really take us much further?

No definition of ratio decidendi

We must in fact admit failure: no one has yet adequately defined *ratio decidendi*. A judge looks for the principle of law as applied to facts that appears to him to be appropriate, and takes that as the *ratio*, and we can be no more precise than that. We can close our discussion of *ratio* by looking briefly at why it might be difficult to identify that principle in particular cases; several obvious reasons spring to mind. Judges do not always explain themselves properly; they often give several different reasons for a decision. Sometimes the actual decision may follow as an exception to a field or rule expressly considered in detail (for example, *Hedley Byrne* v *Heller* (1964), where the House of Lords laid down a new rule on negligent mis-statements but decided the case on an exception to the rule, viz the bank's disclaimer), and even sometimes the decision may not seem to follow from the reasoning. In cases with more than one judge, all saying different things, working out the *ratio* can be impossible. In a case from the US Supreme Court, *University of California Medical School* v *Bakke*, the *ratio* is said to be the decision of one of nine judges. This justice, Powell, agreed with four justices on one point, and the other four justices on another. The accepted *ratio* is thus one with which eight of the nine justices would not agree. When you add to these uncertainties the problems of later decisions, choosing one possible *ratio* (as per Stone), and later courts having to decide on a series of cases in this way, the complexities of discovering *ratio decidendi* become apparent!

It is possible that a case will have no ascertainable *ratio* at all. This, according to de Smith, *Constitutional Law*, is the case with *Nissan* v *Attorney-General* (1970) concerning a claim for damages caused by British troops billeted in a Cyprus hotel where the judges in the House of Lords all gave separate reasons for their decision. The case of *Harper* v *National Coal Board* (1974) shows a further difficulty. This was a decision of the House of Lords in which by a majority the decision went one way and the reasons went the other way!

Those parts of a judgment which are not the *ratio* are called *obiter dicta*. These parts – of however high a court or respected a judge – are like the decisions of lower courts, Privy Council (PC), foreign courts, etc: merely persuasive. Some, especially House of Lords, *dicta* are treated as near binding – the statement of principle in *Hedley Byrne* for example, and the CA discussion of precedent rules in *Young* v *Bristol Aeroplane* (1944). Many *dicta* are ignored or expressly contradicted (just as many non-binding cases are not followed).

Flexibility, so far, has entered the *stare decisis* doctrine via authority of court or report, via choice of what is the *ratio* – because it is much in doubt, of course, it almost goes without saying that later judges have flexibility in choosing what it is – and in disregarding or accepting *dicta*. Judges can even avoid a case that is binding on them by a device known as distinguishing; that is, taking it as not covering the facts of the present case. Obviously the choice of *ratio* is important to this process: choosing the relevant facts for the *ratio* at a different level of generality, or suggesting that facts in the previous case which do not appear in the present case were material to the decision. All law students can remember instances of this, and also instances of cases where earlier decisions have been treated as authority only on their own particular facts. In these ways, judges can distinguish past cases, and limit their precedent effect.

The doctrine of *stare decisis* appears fixed and settled; in practice it is a flexible weapon in the hands of a judge. A core area of fixed law is surrounded by a fringe area in which judges, by distinguishing, approving and following past cases, steadily develop the law.

24.4 The rules of precedent

It would be appropriate for general background information to include here a summary of the rules of precedent as they apply in each of the main courts in this jurisdiction.

The House of Lords

The *Practice Statement* adopted by the House in 1966 changed the previous practice of the House of Lords, laid down in *London Tramways Co* v *London County Council* (1898). The previous rule was that the House of Lords would not depart from its previous decisions under any circumstances; the 1966 *Practice Statement* stated that they would do in future if it was right to do so. Their Lordships remained aware of the importance of certainty in the law (particularly in relation to contractual etc arrangements and criminal law), but strict obedience to past decisions could cause injustice and restrict development of the law.

Some surprise has been voiced that this change was made in a Practice Direction. However, rules of precedent do not form part of the *ratio* (nor do rules of statutory interpretation, see below) of cases, and are just judicial practice.

A more interesting question is whether the House of Lords should have changed the rules. It seems to me that the *Practice Statement* was a good thing, allowing the House of Lords to be honest in their treatment of past authorities now felt to be unsatisfactory. Rather than distinguishing, they can now overrule. Certainty is a virtue, but one that can be over-indulged in.

The *Practice Statement* has been directly used less than a dozen times in the last twenty-two years: *Miliangos* v *Frank* (1976), overruling *Re United Railways of Havana and Regla Warehouses* (1961); and *ex parte Khera and Khawaja* (1984), overruling *ex parte Zamir* (1980), are two of the examples.

Is the Court of Appeal bound by the House of Lords?

Recently the Court of Appeal has attempted to free itself of House of Lords' dominance in relation to House of Lords' cases it finds unacceptable. In *Cassell* v *Broome* (1972), the Court of Appeal said that the House of Lords' decision in *Rookes* v *Barnard* (1964) was arrived at *per incuriam* (that is, without citation of relevant binding authority, in this case two previous House of Lords' decisions). The decision was inspired by Lord Denning and was the subject of almost unjudicial condemnation in the House of Lords when the matter went on appeal.

In *Schorsch-Meier* v *Hennin* (1975) the Court of Appeal refused to follow the House of Lords' decision in *Havana Railways* that currency judgments must be expressed in sterling, on the basis that the reason for the rule had gone: *cessante ratione legis: cessat ipsa lex* (if the reason for the rule ceases, so does the law). In *Miliangos* v *Frank* the House of Lords deplored the Court of Appeal action in *Schorsch*. Strict adherence to the hierarchy was required for the precedent system to work.

The Court of Appeal is bound, then, to follow the House of Lords loyally. The problems when it does not, as in *Schorsch*, can be seen from the dilemma of the first instance judge, Bristow J, in *Miliangos*. Should he follow the House of Lords' decision, or the later (but heretical) Court of Appeal? In fact he followed the House of Lords (the Court of Appeal followed itself in *Schorsch*). Opinion is divided on whether Bristow took the right side, but is united on the difficulty of his position!

Is the Court of Appeal bound by its own past decisions?

In *Young* v *Bristol Aeroplane Co Ltd* (1944), Lord Greene MR laid down the still applicable position for the Court of Appeal. It is bound by a past Court of Appeal decision, unless:

1. There are two conflicting decisions – one must be overruled.
2. While not expressly overruled by, it is nonetheless inconsistent with, a subsequent House of Lords' decision.
3. It was arrived at *per incuriam* (relevant binding authority not cited).

We should note two other exceptions: the Court of Appeal is free to follow a later Privy Council decision inconsistent with a previous Court of Appeal decision: and in a criminal case, the Court of Appeal is not bound if it would cause injustice in the instant case. (Remember Court of Appeal (Criminal Division) cases do not bind Court of Appeal (Civil Division) and vice versa.)

In recent years, the Court of Appeal led by Lord Denning has shown an anxiety to throw off these shackles. Lord Denning has said that the Court of Appeal is not bound by previous decisions (*Barrington* v *Lee* (1972) for instance) and that the Court of Appeal could issue a Practice Statement similar to the House of Lords (*Gallie* v *Lee* (1971)). He did not always carry the Court of Appeal with him, but he did lead a five-man Court of Appeal in *Davis* v *Johnson* (1979) which purported to overrule two Court of Appeal cases (*B* v *B* (1978) and *Cantliff* v *Jenkins* (1978)) on the Domestic Violence and Matrimonial Proceedings Act 1976.

In *Davis* Lord Denning said the Court of Appeal should either follow the direction of the House of Lords *Practice Statement*, or add exceptions to *Young* where appropriate. Both come to the same thing: he was claiming that the Court of Appeal could overrule its own previous rulings. The other two in the majority, Baker and Shaw LJ, drew up new exceptions to add to *Young*. The House of Lords roundly condemned the Court of Appeal, reaffirming *Young*. (They did however overrule *B* v *B* and *Cantliff*.)

Should the Court of Appeal be bound by its previous decisions? Bearing in mind that the House of Lords changed the no overruling ourselves rule by a Practice Direction, and that the Court of Appeal can arrange its own procedure, can it issue a Practice Statement on the same lines as Lord Gardiner's in 1966?

An alternative: prospective overruling

The main argument for *stare decisis* is certainty. Certainty is a value in a legal system because it allows people to arrange their affairs in accordance with the law, both not breaking it (crime) and taking advantage of its facilities (contract, wills, etc). If judges departed from their decisions at will, these arrangements would be upset; further, the individual case would be in effect a retrospective law, changing the law as it was and applying the new law to the present case.

In the case of *Great Northern Railway Co* v *Sunburst Oil* (1932), a decision of the United States courts, Cardozo J stated that in order to avoid this problem the court could adopt prospective overruling. This is a method of treating the present case on the old law, but announcing the new law for future cases. This only, of course, avoids the retrospective argument; could it be so arranged (for example, by applying the new law to future arrangements only?) to avoid affecting settled arrangements? Also, would it not be extremely unfair to the losing litigant, who would have persuaded the judge(s) to accept his legal argument but still have lost the case?

The question that is really being asked is whether certainty and development of the law go together?

24.5 An introduction to statutory construction

A subjective approach

More perhaps than is the case in most of the Chapters in this textbook, the selection of topics and contents for this Chapter, and the arrangement of them, is a very subjective one. A quick glance through any of the major textbooks on either Jurisprudence or English Legal System will show that each approaches this area differently, emphasising different points and using different case illustrations. Most of these textbooks would agree, though, in recommending Professor Cross' *Statutory Interpretation* (1976) to any student of the subject.

Interpretation

As our law becomes increasingly statutory, with upwards of 60 public Acts of Parliament each year (as well as innumerable statutory instruments), the interpretation of those statutes becomes increasingly the judges' central role. There will always be a need for such interpretation and construction. Words are ambiguous, phrases and paragraphs are more so; and no legislator can cover every possible future case clearly. Since under our constitution matters of law are decided by the judges, the task of working out the meaning of the unclear statutory provision, and seeing if it applies to the (frequently unforeseen) case before them, falls to the judges.

Ambiguity

Various sorts of problems can arise. A distinction is often attempted between interpretation (deciding the meaning of the words) and construction (seeing if the words apply to a particular case): the definitions in brackets are only one variant. I will not use this distinction, but will instead bear in mind that, apart from those cases where the meaning is obvious and straightforward (enabling both the judge and the layman organising his affairs to see what the statute means immediately), there are cases where a particular word or phrase is ambiguous, cases where it is unclear whether a particular fact-situation was meant to be included, cases where the particular punishment intended is not clear, cases where the legislature appears to have left out an obvious case, and cases where the result on the straightforward meaning of the words is absurd.

Intention

Note how often I have used the word meant. Judges often say that they seek the intention of Parliament: the great debate between the literal meaning and the mischief-purpose approach is said to hinge on whether Parliament's intention is to be gleaned merely from its exact words (he meant what he said) or also from a

consideration of why the statute was passed (its purpose) and what Parliament would have done if it had had the particular case in mind. Any search for intention, purpose, etc, is to an extent a fiction. A body like Parliament is made up of many people, who may not vote at all on a measure, or may vote for the measure for tactical reasons without considering its consequences, or may vote for it for tactical reasons apart from the actual content. Often votes are on general principles, and yet the matters that come before the courts will be detailed and perhaps highly technical.

To that extent, then, one cannot say what Parliament intended. However, the judges are looking at Parliament's words and must (under the Parliamentary supremacy doctrine) follow and attempt to apply those words. While guidance may not be available on a particular matter, it is clearly the case that, on general principles at least, it does not seem so absurd to search for a Parliamentary intention. Surely the Sex Discrimination Act was *intended* to remedy some aspects of discrimination against women, the Unfair Contract Terms Act was *intended* to control exemption clauses and the Supplementary Benefits Acts are *intended* to set up a scheme providing those with no income with a state safety-net? And more specific provisions can be seen to be *intended* – a provision repealing an earlier provision or overturning an earlier case; a provision following a Law Commission recommendation where no one in Parliament argued with the Commission's reasons. Whether it be intention of the draftsman, or intention of the proposer, or intention of the majority, there is some sense in the concept of Parliament's intention.

Having said that, again I emphasise that most often in difficult cases Parliament's intention is not clear. On a disputed provision, did Parliament intend to protect from that specific type of exemption clause? It is precisely because the words do not make clear what the intention is that the problem arises in that case, and in general, if the words are not clear, how are the courts to decide what Parliament's intention was? To put it another way, what do the words as enacted by Parliament legally mean?

Statutory interpretation

The final introductory part concerns the status of decisions on the question of statutory interpretation. Assume that the House of Lords has to deal with statutory provision X1; the plaintiff claims it means X2 and the defendant X3. The House uses the literal method, and finds for the plaintiff. What is binding on lower courts? Clearly, it would seem, not the literal approach; the rules of statutory construction do not appear to be part of the *ratio* of any case; surely it is only the decision that, in this statute, X1 = X2. If the same words occur in a different statute, the different context and purpose might justify a different result; but on the same statute, lower courts would be bound to follow the House of Lords.

24.6 Canons and presumptions

Apart from the major rules considered here, in cases where statutory words are obscure or unclear judges may use one or other of the following canons of construction and presumptions.

Canons

The statute must be read as a whole
The words of the particular sub-section in question must not be read in isolation, but must be read with the other sections (particularly any interpretation section) and with the schedules. As we shall see below, this canon is now subsumed by Professor Cross' reformulation of the major rules, where he emphasises that the context of the words is in account.

Eiusdem generis
If a general word follows two or more specific words, the general word must be restricted in meaning to a meaning of the same kind (*eiusdem generis*). For example, *Powell* v *Kempton Park Racecourse Co* (1899) turned on whether in relation to places of betting the words house, office, room or other place included the racecourse itself: No, said the House of Lords, since the general words other place were restricted to a meaning of the same kind as the specific words, that is, an indoor place of betting.

Narrow construction of penal provisions
The individual gets the benefit of any doubt if a criminal or tax liability is imposed by statute, in particular against the imposition of liability without fault.

Interpretation Act 1978
This Act gives presumptive interpretations to common words and phrases in statutes: so men includes women (and vice versa), singular includes the plural, distances are to be measured in a straight line on the horizontal plane, time refers to Greenwich Mean Time and so on: all subject to contrary intention (which must sometimes be expressly stated, but most often must just appear).

Presumptions

Against alteration of the law
This presumption does not work against a change in the general (common) law which appears clearly from the literal meaning of the words; but if there is a doubt, Parliament will be presumed to have left the law unaltered.

Against imposition of without-fault liability
Mentioned above; to create a strict liability offence, Parliament must use clear words.

Against ousting the jurisdiction of the courts
The courts are very protective of their own jurisdiction; although Parliament may alter the courts' jurisdiction even fundamentally, it must do so clearly. In administrative law, for example, in several cases the courts have evaded statutory attempts to forestall judicial review (*Anisminic* v *FCC* (1969), *Padfield* v *Minister of Agriculture* (1968), *Pyx Granite Ltd* v *MHLG* (1960)).

Against the Crown being bound by a statute
The Crown must be expressly named, or it is not bound by a statute.

Against depriving a person of a vested right
The above are just examples. It may be quite possible to find canons and presumptions to support quite conflicting contentions.

24.7 The three rules of statutory construction

It is often said that there are three rules of statutory interpretation, these being the literal, golden and mischief rules. As we shall see they are to an extent contradictory; all can claim judicial support.

Mischief rule

This rule was prevalent in the sixteenth century. The courts have regard to the purpose of the Act, and interpret it in such a way that the purpose is fulfilled or enhanced. The classic statement of the rule is contained in *Heydon's Case* where the barons laid down four things to be considered when interpreting statutes: the common law before the Act, the mischief that the law did not provide for, the remedy appointed for that mischief, and the true reason of the remedy. Of course, not all statutes are altering the common law today, and the exact formulation therefore needs changing. The approach, while not now as prevalent as it was, still commands judicial support, and has authorities following it in many areas (see the examples given in Dias' *Jurisprudence*).

A recent example can be taken from the law against racial discrimination. Although there is a requirement in the mischief rule that the express words of the statute must reasonably bear the purposive meaning given to them, in *Mandla* v *Dowell Lee* (1983) the House of Lords interpreted the Race Relations Act, where it is stated that it is an offence to discriminate in certain matters against a person on grounds of his race, colour, ethnic or national origin, in quite a different manner.

The facts of the case were that a young Sikh male wanted to join a public (fee paying) school. He was granted admission but was required to conform to uniform regulations and remove his turban and cut his hair. For reasons of faith he was unwilling to do this. There were other Sikhs in the school who had conformed to the uniform requirement and there was no suggestion that Dowell Lee (the headmaster) had any inclination to discriminate against Sikhs. The Court of Appeal carefully considered the history of the Sikh people and concluded that they were a group identifiable only by their common religion and that as the statute makes no mention of religion then the actions of the school were reasonable and not illegally discriminatory. The House of Lords, relying on a New Zealand case concerning the position of the Jews (*King Ansell* v *The Police* (1974)), held that the purpose of the section was to cover situations such as the present and that by a stretch the Sikhs could be regarded as a group identifiable by a common ethnic origin. The reason for so holding was to extend the protection afforded by the Act to Sikh people.

Had the court been minded to find otherwise then it might have followed the case of *RRB* v *London Borough of Ealing* (1972) which held to the literal approach (see below) in holding that discrimination against a Polish citizen in the granting of public housing was lawful because it was not on grounds of his national origin but on grounds of his citizenship or nationality. Perhaps this comparison between these two cases reinforces the view that in their choice of which rule of statutory construction to apply the judges in effect determine the outcome of the case. Bishop Hoadley put it thus centuries ago: 'Whoever hath an absolute authority to interpret any written or spoken laws, it is he who is truly the lawgiver to all intents and purposes, and not the person who wrote or spoke them.' Lord Devlin perhaps has it better and in more modern language when in his *Samples of Lawmaking* he states that 'the law is what the judges say it is'.

The mischief or purposive rule is the one favoured by Fuller as elaborated upon in his 'The Case of the Speluncean Explorers' which I discussed in the third Chapter of this book. Perhaps the reader would return to that article at the end of our course and extract more from it. Before you do that let us now examine the other rules:

Literal rule

Various factors, including the declining influence of the judges on legislation and the development of Parliamentary supremacy, led to a retreat from the mischief type approach to the literal approach. Here, the intention of Parliament is considered as contained in the words passed: the literal meaning of those words must be taken, even if the result appears to be one which Parliament did not intend. Lord Esher in *R* v *City of London Court Judge* (1892) stated that 'the court has nothing to do with the question whether the legislature has committed an absurdity'. This follows on the constitutional provision that it is the role of the legislature to make law and the role of the judiciary to interpret the law the legislature so makes.

Many cases support this rule of applying the clear and unambiguous words of

Parliament. For example, in *Inland Revenue* v *Hinchy* (1960) the House of Lords was construing a provision which visited upon people incorrectly completing tax returns a penalty of treble the tax that ought to be charged under this Act. Presumably Parliament intended the punishment to be three times the excess owed: but those words meant three times the whole tax bill for the year, which cost poor Mr Hinchy £418 instead of £42!

Note at this stage two things. First, words are often not clear and unambiguous; two equally usual meanings of a word might exist, or the application of words to particular cases might be in doubt, and so on. Second, it is not unknown for judges to consider the literal meaning of the words and end up with different results (for example, *Liversidge* v *Anderson* (1942)).

Golden rule

Judges have often mitigated the strict literal approach by calling into play the golden rule, that is that if the usual interpretation results in consequences so absurd that Parliament could not possibly have intended them, any secondary meaning may be taken. In the case of *R* v *Allen* (1872) which concerned the definition given to the offence of bigamy in the Offences Against the Person Act 1861 as 'whoever being married, marries another' where it was observed that such a definition if applied literally would lead to the absurd conclusion that the offence could never be committed. A person cannot legally marry he is are already married. There the court held that, as Parliament could not have intended to legislate nonsense, the words should be changed to read 'whoever being married goes through a marriage ceremony with the intention to marry etc'. Then the definition has meaning which would probably be consistent with the intention of the legislature.

Obviously, the three rules above cannot really be taken as strict rules: they contradict each other (taking the literal meaning often obscures the purpose of the statute, it might be said). At most they are approaches, with the judges choosing the most appropriate in the circumstances, generally plumping for the literal rule and taking the obvious plain meaning unless some good reason to the contrary appears.

Even this does not seem to be a good explanation of what happens if we accept that the judges generally follow the approach of looking at the literal meaning. What of those cases where two meanings are equally usual and neither of the other two approaches is relevant or helpful? What of technical words?

A rather more successful attempt at formulating the courts' approach overcoming the lack of judges giving reasons has been made by Professor Cross in *Statutory Interpretation*. He suggests that the literal and mischief rules have been mixed, and the vital element of context added: the judges look to see what the ordinary (or, if appropriate, technical) meaning of the words used is in the general context (including the objects) of the statute. It is that ordinary meaning that may be displaced by a secondary meaning if the result would otherwise be absurd: and furthermore, in cases where what seem like simple mistakes make a statute

unintelligible, absurd or totally unworkable, a judge may add or delete words, to change nonsense into sense (Cross cites *Adler* v *George* (1964) and Lord Denning in *Eddis* v *Chief Constable* (1969)).

The whole problem stems from the Blackstonian fiction that statutes are intended to govern all eventualities in detail and do not merely lay down guidelines. Taken with the imprecision of words – a problem Hart has dwelt upon when he referred to the core of settled meanings and the penumbral area of doubt that surrounds words – the problem of statutory construction is manifest. This is clearly stated by Lord MacDermott thus: 'the difficulty of finding unequivocal language by which to convey the will of Parliament [lies at the heart of the problem of statutory construction]'.

The Swiss are perhaps more realistic. Their Civil Code in Article 1 states that a judge may decide a case on the basis of a rule which he would lay down if he himself has to act as legislator. The only limitation in this regard is contained in Article 4, to the effect that in exercising his discretion the judge must base his decision on principles of justice and equity. Lloyd and Freeman observe that although this article was initially widely used it is now subject to restrictive interpretation itself. Does this tell us something about the nature of the judicial creature?

Hypothetical examples

Much of what judges do is obvious, even when they construe difficult or ambiguous sentences or phrases: although we must consider the pros and cons of judges following a purposive as against the traditional literal approach, we must also emphasise that in fact it is in comparatively few cases that a straightforward literal v purposive clash occurs. The following fact situations might help to make the point.

1. A particular word or phrase has a straightforward obvious usual meaning, for example driving a motor-car at over 70 mph is an offence. A driver knows that once the speedometer tops 70 he is committing an offence, the judge when he is deciding applies the obvious meaning of motor-car, driving and 70 mph and convicts. This is straightforward literal approach: in relation to this case, the words have only one meaning.
2. A particular word or phrase has several meanings: for example, the verb wants (wishes or lacks?) the noun will, (volition or the document by which a deceased person leaves his property?).

   The context of the phrase in the statute makes it clear which sense is meant: for example, a reference to providing what a lunatic wants will refer to what he lacks; a reference to the will of the testator in a statute on probate will generally mean the document (but could in context mean the volition, as, 'the will of the testator was overborne by force').

   The judge applies that obvious meaning. Not quite the literal approach, since

there were two usual meanings (and in the case of wants, the one chosen was, if anything, the less obvious or usual of the two). But can this really be called a purposive approach? We are looking at the in-context meaning, and purpose is relevant only as part of the context.

3. As situation (1) except that this meaning either produces an absurd result, for example, (ignoring the Interpretation Act) it is an offence to steal horses, and the defendant steals just one (so he is not guilty under literal meaning), or produces a result clearly against the intention of the Act, for example, if the Race Relations Act defined racial group in a technical way which excluded negroes.

   As to the absurd result, holding that the statute meant something else, this clearly involves the judge in rectification, which Cross allows as his third rule; not the golden rule, as there is only one meaning the words can bear (and therefore no secondary one to fall back on).

   As to the result clearly against the intention of the Act, any suggestion that the judge acts in accordance with that intention and not the words of the Act does lead to a purpose v intention conflict. Note, however, that in general the courts have not invoked the mischief rule in this sort of case: an attempt by Lord Denning to fill the gap left in a statute in the case of *Asher* v *Seaford Court Estates* (1949) was slapped down by the House of Lords, Viscount Simmonds rejecting this naked usurpation of the legislative role *(Magor & St Mellons RDC* v *Newport Corporation* (1951)). If the result is not absurd, the courts will follow the wording of a statute if it only allows of one construction, even if that construction does not follow the general purpose of the statute.

4. As situation (3) except that one meaning is clearly the more usual, but that result leads to either absurd consequences or is totally against the intention of the statute. An example of absurd consequences could be the facts of the tax case *Inland Revenue* v *Hinchy*. An example of being against the intention of the statute can be seen in the United States controversy over whether reverse or positive discrimination is against the constitutional provision; forbidding laws which deny equal protection of the laws: does that mean that any discrimination is unlawful or could 'equal protection' be taken to include the effect of reverse discrimination in redressing the balance and hence making more equal?

   If the judge takes a secondary meaning to avoid absurdity, that is the golden rule in operation; if he takes it to accord with the intention of the statute, that could be taken as using the context of the statute, if not (and in our example, the context doesn't help: the question is, how far did the constitution go?) he is using purpose to displace the literal rule.

5. As situation (2) except that the context does not assist, the purpose of the statute does not assist, and the consequences would not be (more) absurd either way. The judge uses his discretion – but none of our stated approaches/rules!

Literal words v *purpose*

In (3) and (4) then, there are possibilities for a clash between words and purpose: should the judge follow the obvious or only meaning of a phrase or sentence if that goes against the purpose of the statute? Briefly, the arguments for the literal approach are: certainty; avoidance of judicial legislation; due deference shown to Parliament; it is often difficult to identify purpose; and it encourages more careful drafting. For the purposive approach: it is often not possible to work out what the literal meaning is; it is not really deferent to Parliament to refuse to fulfil its purpose; and judicial legislation is common, particularly in the common law.

Which of these sets of arguments convince you? Are there any other points to be made?

24.8 Aids to construction

Where a statute's construction is ambiguous or uncertain, various aids may be used by the judge to help him come to his decision (to minimise tedium, case references are omitted).

The rest of the statute

A statute must be read as a whole, as we have said above; the judge must therefore decide in the light of the rest of the enactment (including the long title). In cases of uncertainty, those parts of the statute which are not integral parts of it (preamble, marginal notes, punctuation) may be called in aid.

Other statutes in pari materia

If construction is uncertain, a statute on the same subject may be called in aid, if it is unambiguous.

International treaties

If an Act is stated to be intended to give effect to an international treaty, uncertainties may be decided by reference to the treaty.

Statutory materials: The landmark case of Pepper *v* Hart

Pepper (Inspector of Taxes) v *Hart* (1993) is an important case on statutory interpretation. It settles, for the time being, the question of the extent to which 'extrinsic' materials may be used in ascertaining the extent of that well-known personification of the *point* of legislation, parliamentary intention. The House of Lords, in which seven judges appeared, came down in favour of a relaxed rule

(although the Lord Chancellor, Lord Mackay, dissented). It was thought that the use of 'parliamentary materials' would be permitted in legal argument regarding the discovering of the purpose of legislation in the following cases:

1. where the legislation was ambiguous or obscure, or the literal meaning led to an absurdity;
2. where the material relied on consisted of statements by a minister or other promoter of the Bill which led to the enactment of the legislation together, if necessary, with such other parliamentary material as was necessary to understand such statements and their effect; and
3. the statements relied on were clear.

Furthermore, the use of parliamentary material as a guide to the construction of ambiguous legislation would not infringe s1, article 9 of the Bill of Rights since it would not amount to a 'questioning' of the freedom of speech, or parliamentary debate. This was provided counsel and the judge refrained from impugning or criticising the minister's statements or his reasoning, since the purpose of the courts in referring to parliamentary material would be to give effect to, rather than thwart through ignorance, the intentions of Parliament. Furthermore, the use of parliamentary material in this way would not question the processes by which such legislation was enacted or criticise anything said by anyone in Parliament in the course of enacting it.

Pepper v *Hart*, in effect, follows the recommendation of the *Renton Committee on the Preparation of Legislation* (1975) that the courts should accept constructions promoting the general underlying purpose.

24.9 Effect on the draftsmen

Past and present practice of the courts on statutory interpretation clearly affects how draftsmen work on future legislation. An example from the Wills Act 1837, cited by Cross *(Statutory Interpretation*, p12), shows how ridiculous were the lengths to which draftsmen then were driven to avoid the rigours of the full-blown literal approach. The courts are not quite as exacting any more, and do take at least the context into account with the words, but the enduring pre-eminence of the literal approach and the eagle eyes of eager lawyers intent on taking every possible point for their clients do still affect the form and structure of present legislation.

Procedure

Generally, the procedure for drafting is a careful one, especially if the statute is lawyers' law, rather than that dictated by party policy. For example, the Law Commission will issue a working paper, followed by a report with draft Bill, or the government will issue draft proposals (in Green or White Paper form) for

consultation. As much time as possible is given to allow lawyers and others to look for, *inter alia*, drafting mistakes.

Detail

Often statutes go into great detail to avoid unwanted interstitial interpretation: for example, Employment Act 1980, in its sections defining the outlawed secondary action and secondary picketing.

Examples

Many statutes give examples of the instances intended to be covered as the factors to be taken into account: for example, 1973 Matrimonial Causes Act, ss23–25, detailing the factors to be taken into account by a judge in deciding the financial provision on divorce as examples (because all the circumstances are in account).

Discretion

When judges are intended to have discretion on a particular matter to decide in accordance with the statute's purposes, this is sometimes expressly stated in terms. Section 23 Matrimonial Causes Act is again a good example; the judge must do what is just and equitable in all the circumstances in an attempt to put the parties in the position they would have been in if the marriage had not broken down.

Interpretation

Many statutes contain their own interpretation sections.

24.10 The common law and the Constitution

An exciting, radical approach to judicial decision-making was advocated by Laws J in *Public Law* (1995). His thesis sent shock waves amongst public lawyers and will clearly have an effect on the future direction of the increasingly significant use of the procedure of judicial review. It is most convincing and its chief significance lies in its public avowal by one of the most talented judges in the United Kingdom that judges, by virtue of their inherent jurisdiction to declare whether something is required or permitted by law, could declare purported statutes to be legally invalid. This, of course, cuts through an (unanalysed) general assumption that Crown-in-Parliament may make any laws that it pleases and that it would be wrong for a judge to 'usurp the function of the legislature'. The reasoning of Laws J is as follows. Inherent in the idea of the legitimacy of Parliament is the idea of democracy; that idea, whatever else it means, stands for each individual's stake in the legal system.

'One man, one vote' must mean at least that. What would it mean to give each person a vote and yet at the same time deny that each person had any right to expect a certain minimal level of treatment? Or, abstractly, the idea of democracy entails that people, by virtue of being people, are entitled to a certain level of respect. We do not deny that Parliament has the right to make laws for us by virtue of the fact, amongst others, that it is elected by us. If that is so, if Parliament does something *contrary to the principles that give it meaning* as a legislative body, it is acting *ultra vires*.

This idea is most attractive idea because it gives weight to our intuition that (*contra* Austin) there is more to the idea of legislation than brute power – the mere ability to enforce a command. The form of argument used here is a 'transcendental' one; you look to the nature of Crown-in-Parliament and then you deduce from that nature some other principle which transcends the idea and then acts as a constraint upon it. (Remember the 'transcendental epistemological deduction' of the Grundnorm in Kelsen – see Chapter 9, above) This sounds more difficult than it really is but becomes clearer by using a fairly unlikely – but nevertheless still realistic – example. What if Parliament decided to abolish the vote for unemployed people (perhaps out of a misguided sense that since such people did not work, they did not contribute to society and thus had no 'stake' in our community)? Wouldn't that be contrary to the very principles (of democracy) that lie at the heart of our legal system? What obvious and compelling reasons would there be for denying a judge the right to declare this legislation invalid because ultra vires?

It is useful to array the arguments on either side:

For

1. It makes sense using the transcendental argument referred to above; from what other principles could Crown-in-Parliament gain its validity?
2. It is clear that in other jurisdictions there is no problem in the idea *at all*; the United States, example. To the reply that there is a written constitution there but not here we could add;
3. There is no significance in the distinction between a written and unwritten constitution in the same way that there is no significance in the distinction between the telling of a story and the writing down of a story, as far as the story itself is concerned. To the reply that writing it down makes the story more certain we could add;
4. There are just as many problems with the interpretation of the written word as there are with the spoken word (witness the enormous litigation in the United States which is still created by the famous phrase 'equal protection of the laws' in the constitution).

Against

1. Very few people in the United Kingdom would accept the idea; in particular, one of the most influential textbooks, *Wade on Administrative Law*, is firmly of the view that Crown-in-Parliament can enact whatever it wants (and this is, of course, an idea as old, even older, than Austin).
2. If Laws J is right, it means that Parliament could never repeal, for example, the relevant parts of the 1832 Reform Act, which greatly extended the franchise.
3. We could pay attention to Bentham's idea of a 'limited disposition to obey'. Bentham, unlike Austin, did not think that the sovereign was legally limitless, and he explained his belief by saying that the sovereign depended upon the ability to command by the willingness to obey displayed by the population at large. To take an extreme example: if Parliament passed a law, similar to the one passed in ancient Sparta, which declared that all male babies were to be kept outside unclothed all night so that only the fittest would survive, *no one* would obey it. It followed for Bentham that it would not be a law, despite the fact that the sovereign had 'commanded' it (in reality it is a failed command, because the threat is not real given the command's content), and so Bentham can preserve his distinction between the 'is' and the 'ought' of law by burying away the idea of the legitimacy of reasons for obeying the law inside the idea of a command.
4. Judges should not make 'political' decisions. But, as Laws J points out, it all depends upon what is meant by 'political'. Judges are political in the sense that their decisions have political consequences; the legislature, true, has certain sorts of competence carved out for it (a judge could not make fundamental decisions about the direction of the economy, for instance). Nevertheless, the legislature cannot be immune from legal criticism for acting outside the principles of democracy fundamental to its legal legitimacy (it could not further economic policy by killing off the unemployed, for instance). Thus Laws J says:

> '... the *subject-matter* of a case offers no inhibition to legal adjudication on grounds of its political content.'

This article will have far-reaching effects since it strikes at the heart of commonly accepted (but confused) propositions about the immutability of the principles of statutory interpretation in the light of 'the intention of Parliament'.

25

Dworkin's Law as Integrity

25.1 Introduction

Dworkin's theory is fascinating and highly practical. He is difficult to get into because his writing output is enormous and many articles are difficult to obtain. He is the most important contemporary in legal philosophy. One of the difficulties is that, although he is in the rigorous intellectual mould of Bentham, Kelsen, Hart and so on, he is *not* a positivist. The best start is to read Chapter 2 of his *Taking Rights Seriously* (1978). Students should then go on to read Chapter 4 and then read his *Law's Empire* (1986) especially Chapters 2, 4, 5, 6 and 7. The important point with Dworkin is not to underestimate his subtlety and intellectual power.

25.2 Hercules, the model judge to whom we should aspire

Let us get clear about Hercules. Many people, in particular lawyers, who are introduced to Hercules in Dworkin's article 'Hard Cases' simply dismiss him by saying that no such judge ever existed. This is too glib. Hercules is a model against which, like any other ideal, legal arguments are to judged. Take, for example, the idea of the ideal market. It would be off the point to say that the ideal market does not exist (or that the model of the atom, or that of the DNA molecule, does not exist). To *say* that, is to recognise the idea of the ideal, in any case. The point of the ideal market, about which economists and practical minded politicians argue

vociferously, is to show how, in the real world, there are imperfections. Against the model of the ideal market, we see that monopolies and other restrictive practices are 'bad', that transaction costs and imperfect knowledge 'distort' the real market, and so on.

Why not, then, imagine the existence of an ideal judge, against which we can measure bad or distorted legal arguments? There is no reason to suppose that we cannot. But it is worth trying to explore the reasons why people make the mistake of saying, in effect, that there cannot be ideal arguments in the law. The problem is one of superficiality, no more. People like to think of law as historic fact. They do not like to think of legal argument as something as shifting and as controversial as moral argument.

It is necessary for Dworkin to posit an ideal judge because his theory is about law as an argumentative attitude (see Chapter 3). He has to provide a scheme of argument which, among other things, is sufficiently abstract to allow for controversial argument. He cannot provide a set of premises from which conclusions may be drawn by, say, the use of syllogisms. His is not that sort of theory. In fact, he is critical of that sort of theory. He thinks it paints a simple-minded picture of legal reasoning. In order for him to describe the inherently controversial nature of hard cases, he can only provide the general scheme of argument.

Dworkin's device of Hercules is used to characterise correct legal argument. It is not that there is a method which will come up with the right answer, there, uncontroversially for all of us to see. If a problem is raised about whether there *could* be such a right answer, it is one about the objectivity of legal argument, not a criticism of the ideal model of Hercules.

Students should note that a new collection of essays by Ronald Dworkin will be published in this country shortly, by Oxford University Press. It is already published in the US as *Freedom's Law: the Moral Reading of the American Constitution* (1996). It is divided into three sections entitled Life, Death and Race; Speech, Conscience and Sex; and Judges. Many of the articles have appeared over the last ten years in the *New York Review of Books*. Students should read his Introduction: the Moral Reading of the Constitution, as it is particularly helpful for obtaining yet another angle on how he thinks judges should decide cases.

25.3 Hercules and 'hard cases'

The key to understanding how Dworkin thinks a hard case should be decided is in the following idea:

> 'If a judge accepts the settled practice of his legal system – if he accepts, that is, the autonomy provided by its distinct constitutive and regulative rules – then he must, according to the doctrine of political responsibility, accept some general political theory that justifies these practices.'

Look at the position of the judge. He has convictions about his role, his duties as defined by his judicial oath and by other sources. If he did not, it would certainly be

surprising. He has an idea about legislative purpose and principles of the common law.

In this, United States and United Kingdom judges differ. In the United States, judges are more aware of their role as public 'protectors' of the Constitution and they are more explicit (both to themselves and the public) about their position within the separation of powers doctrine and their duties to protect the rights of the individual. United Kingdom judges are, for cultural reasons to do with background and education, more diffident about such matters.

What does Hercules do when constructing the arguments in all the hard cases put before him? We can assume, says Dworkin, that he accepts most of the settled rules of his jurisdiction, rules which lay out for us what are the familiar characteristics of the law. For example, the constitutive and regulative rules that grant the legislature the powers of legislation give judges the powers of adjudication and the duties to follow previous cases, as well as all the settled rules of the various areas of law, such as tort, contract and so on.

We can examine all the possibilities just mentioned in the law reports and academic writings. And these are not exhaustive. We can develop our own theories or, if we prefer, our own arguments. Dworkin's use of Hercules is intended to show the general form – the scheme – of the types of arguments that may be used. We can imagine Hercules producing all the theories, with their attendant sub-theories, for all areas of the law. In each topic, he will have to justify the particular settled rules with the substantive theories he has devised.

He will also have to do more. The division by topic will itself be a matter for justification, which will proceed by way of looking to the settled rules for topic demarcation (say, the division between tortious and contractual liability) and devising a theory which explains that division. He might decide that, for some special cases, the importance of the division may be outweighed, as Lord Atkin, but not Lord Buckmaster, thought it was in *Donoghue* v *Stevenson*.

25.4 The chain novel: 'fit' and 'substance'

Legal argument, for Dworkin, in most hard cases, will develop as the result of a tension between two dimensions of argument, one that argues towards a 'fit' with what is accepted as 'settled' law, the other that argues towards substantive issues of political morality. While the twin abstract injunctions in Dworkin 'to make the best sense' of law, and 'to treat people as equals' propels his legal and political philosophy, it is the distinction between 'substance' and 'fit' that forms the cutting edge, for him, of legal argument.

As a preliminary to getting into the idea, we may employ Dworkin's idea of the chain novel. A number of novelists agree to write a Chapter each of a proposed novel. The first Chapter is written by one, the second by another, the third by another, and so on. We can see that there will be certain constraints of 'fit' upon the

author of the second Chapter, and even more on the author of the third Chapter, and so on. Many more lines of fit could be proposed. Is there fit with style, descriptiveness, thematic material, dialect and so on? The important point is that if certain things are accepted as settled within the text of the first Chapter, later creativity is constrained by that acceptance, in order for the other Chapters to be properly part of the novel.

There are familiar responses to this description of 'fit'. It is a matter of argument (or 'opinion'), people say, as to what constitutes 'fit'. It is too 'wooden' to assert that novels cannot allow for the change of name, sex, century, and geography. But what follows from this? That everything and, therefore, nothing, counts as 'fit'? Of course not. We just do hold some framework assumptions, or constraints, constant while we allow others to vary.

Nevertheless, in Dworkin's terminology, the question of 'fit' is itself an interpretive question. For example, the acceptance of the genre of 'novel' for the chain novel is itself open to interpretation. A second Chapter novelist might, for example, decide that the first Chapter is a political tract about conservatism and best seen as the first Chapter of a political manifesto. This interpretive judgment might constrain the way he continued to write the second Chapter. If he did so, perhaps if the chain novel writing project was a commercial one to produce a radio serial, his contract as a chain novelist would be terminated. But there is no reason holding back the *possibility* of making that interpretive judgment, although, of course, his judgment that the first Chapter was the first Chapter of a political manifesto might be difficult to justify.

25.5 Principles and policies

Dworkin is well known for the distinction he drew between arguments of principle, which are arguments about a person's rights, and arguments of policy, which are arguments about community goals. The distinction is important to Dworkin for a number of reasons. First, it is intended to be largely descriptive of distinctions that in fact are drawn by lawyers. Secondly, it represents for him the line to be drawn between the legitimate jurisdictional activities of judges as required by a properly understood democratic separation of legislative and judicial powers. Thirdly, and most importantly for him, it represents his main assault on the most popularly understood version of the moral theory known as utilitarianism.

It is most important to understand the role of the language in the terms he uses. 'Principle' and 'policy' are terms of art for him. Technically, that means he has stipulated meanings for them. He gives definitions for them as follows in Chapters 2 and 4 of *Taking Rights Seriously* and in *Law's Empire* he accepts these definitions without modification.

> 'I call a "principle" a standard that is to be observed, not because it will advance or secure an economic, political, or social situation deemed desirable, but because it is a requirement

> an economic, political, or social situation deemed desirable, but because it is a requirement of justice or fairness or some other dimension of morality.
> I call a "policy" that kind of standard that sets out a goal to be reached, generally an improvement in some economic, political, or social feature of the community.
> Principles are propositions that describe rights; policies are propositions that describe goals.
> A political right is an individuated political aim. An individual has a right to some opportunity or resource or liberty if it counts in favor of a political decision that the decision is likely to advance or protect the state of affairs in which he enjoys the right, even when no other political aim is served and some political aim is disserved thereby, and counts against that decision that it will retard or endanger that state of affairs, even when some other political aim is thereby served.'

Policy causes difficulties for two different reasons, neither of which strikes at Dworkin's thesis.

1. The distinction is not one of content, but of form. This means that for one person a political state of affairs could be a matter of principle and for another it could be a matter of policy. In this way, the distinction stands clear of differences of political opinion. But it is a distinction of substance, too, in the sense that it requires a strong, separate sense in which principles are not reducible to policies. This is borne out by Dworkin's well-known statement that rights 'trump' utilitarian goals.
2. Critics often point to the fact that it is easy to imagine circumstances in which only goals seem to be important, and where rights are of no consequence. Most people, for example, accept the situation in wartime when civilian rights are suspended. Martial law is accepted as a possible option whenever its imposition warrants the pursuit of the desirable goal of winning a war. The problem is thought to be that if there can be situations where there are justifiably *no* rights, or principles, because of the importance of the goal, what criteria could there possibly be for defining principles independently of goals?

   A special category of emergency is well-described in our moral, political and legal thinking. Martial law is 'martial' law. Its imposition is only justified in wartime, when war is 'raging'. Our concern about the suspension of civilian rights under martial law is characteristically about whether there is a situation which justifies its imposition. Many people, for example, felt that the situation in Romania in 1989, although bad, was not bad enough to justify the Romanian government in imposing martial law. Many felt that it was being imposed, not because it was necessary to preserve the existence of civilian rights, but to protect a particular political system from change, perceived as undesirable.

Some judges are either innovative (in the United States, 'activist') in the chain novel sense to which reference has already been made. Sometimes, rather self-consciously, they will refer to their decisions as decisions of 'policy'. A good example is Lord Denning in the *Spartan Steel* case, referred to by Dworkin in *Taking Rights Seriously*:

> 'At bottom I think the question of recovering economic loss is one of policy. Whenever the courts draw a line to mark out the bounds of *duty*, they do it as a matter of policy so as to limit the responsibility of the defendant. Whenever the courts set bounds to the *damages* recoverable ... they do it a matter of policy so as to limit the liability of the defendant.'

But it is clear from other remarks that he makes that at least some of his reasons for his judgment attach to the particular parties and do not look to future general impact. He thought the matter should be decided on the basis of relationship of the parties. If other relevantly similar bodies were excused liability (in this case by statute), in Lord Denning's view, this was a strong argument for excusing the defendant. The argument was not helped, in other words, by appealing to *a novel way of arguing*.

And yet other judges (Lord Denning fits into all the categories, representing different stages of his career) are blatant. They *do* decide policy, in Dworkin's sense, but disguise it. A good example is *DPP* v *Majewski* (1976), in which the House of Lords interpreted the following words of the Criminal Justice Act 1967 as excluding evidence offered by a defendant as to whether he intended or foresaw a particular result in a criminal case relating to drink or drugs:

> 'A court or jury, in determining whether a person has committed an offence ... shall decide whether he did intend or foresee that result by reference to *all* the evidence ...'

One of the arguments used was that Parliament could not have intended to overturn a clear rule of exclusion in the common law. But this was a *criminal* case and there is an equally clear common rule that criminal statutes should be construed in favour of the defendant. It was argued that s8 was 'only' a rule of evidence, but this was evasive and unconvincing. The sense of the decision was that the judges knew the havoc that would be created by allowing defendants to plead drunkenness as an excuse rather than as only a mitigating circumstance.

25.6 *McLoughlin* v *O'Brian*

We should look at Dworkin's analysis of the case decided in the House of Lords, that of *McLoughlin* v *O'Brian* (1983). This should be instructive as to how he views legal argument. Remember that, in his view, substantial arguments (relating to the right people have to be treated as equals) have to be matched to fit (the already existing case law).

The plaintiff in this case learned that her husband and children were involved in a car accident. She set out for the hospital some miles away, and when she got there she was told her daughter was dead and she saw that her husband and other children were seriously injured. She suffered severe shock and she sued, among others, the driver of the vehicle, whose negligence caused the accident.

How should the case have been decided? It was a hard case, because in all the previous cases, the facts involved people suffering nervous shock almost immediately upon the accident occurring and more or less at its scene. In these cases, the people suffering shock were allowed to recover.

Dworkin suggests how Hercules might decide this case. He says that Hercules might begin by considering the following six possible interpretations of the case law:

1. *Success (for the victim) only where there is physical injury.* But we can rule this out immediately because it does not fit the law of tort. It is clear that damages may be obtained for *nervous* shock.
2. *Success only where the emotional injury occurs at the accident, not later.* But, says Dworkin, this would just draw a morally arbitrary line.
3. *Success only where a practice of awarding someone like Mrs McLoughlin would be economically efficient.* If this were simply a matter of economic *policy*, Dworkin rejects it because it does not respect 'the ambition integrity assumes, the ambition to be a community of principle'. The argument is, in other words, to be rejected because it is one of naked policy, ignoring Mrs McLoughlin's right to be treated as an equal.

   There is, however, an ambiguity which Dworkin says is inherent in the idea that a community should aim at efficiency. It may be that people have a *right* to a certain amount of redistribution under some system which aims at economic efficiency. Dworkin leaves a developed discussion of the idea, (which first appears in a very obscure fashion in his article 'Hard Cases') to Chapter 8 of *Law's Empire*, in a highly compressed and difficult Chapter. But for present purposes, it is true to say that Dworkin rejects the economic pursuit of economic efficiency in this sort of case where it consisted solely of the pursuit of overall (undifferentiated) communal wealth.
4. *Success only where the injury, whether physical or emotional, is the* direct *consequence of the accident.* But this interpretation has to be ruled out because it is contrary to fit: it contradicts the clear case law, where there is a test of foreseeability which limits the liability of the person who causes the accident.
5. *Success only where the injury is foreseeable.*
6. *Success for foreseeable injury, except where an unfair financial burden is placed on the person who causes the accident.* ('Unfair' meaning 'disproportionate to the moral blame for causing the accident').

According to Dworkin, (5) and (6) are the best contenders. (1) and (4) are ruled out because they contradict 'fit'. They simply cannot be made to cohere with the previous legal decisions. (2) is ruled out because it is an interpretation that relies on an arbitrary assertion that people at the scene can recover, those who are not, cannot. (3) is ruled out because it relies on policy, not principle.

Let us now examine interpretations (5) and (6). 'Which story,' he asks, 'shows the community in a better light, all things considered, from the standpoint of political morality?' Suppose that interpretations (5) and (6) equally 'fit' the

precedents. Dworkin says that Hercules should construct two abstract principles. First, that community sympathy towards individuals who are suddenly required to pay large amounts for accidents they cause is an argument in support of public insurance schemes, safety regulations and so on. This is a principle of 'collective sympathy', he says. Second, that it is right that people who are at fault should pay for the consequences of their fault and so costs should be apportioned between private individuals.

These are two principles of exactly the sort that a lawyer could produce in court. If the 'private apportioning' principle should prevail of these two, then interpretation (5) is the correct one. Mrs McLoughlin wins just because Mr O'Brian was at fault. If, on the other hand, the 'collective sympathy' principle prevails, Mrs McLoughlin loses. Why? Because Mrs McLoughlin's injury, while foreseeable, was so remote as to place an unfair burden upon Mr O'Brian in proportion to his fault.

Which interpretation should be preferred? Dworkin thinks that Mrs McLoughlin should have won, favouring interpretation (5), at least in automobile accident cases when there is a widely available and sensible liability insurance obtainable privately.

25.7 The 'one right answer' thesis

Dworkin's point has always been, in line with his theory of interpretation (see Chapter 2, section 2.6), that it makes 'best sense' of our legal practices to suppose that we are all – judges, students, lawyers – striving to argue for, decide, or discover the best answer. That means that there is a 'best' of the matter, and this is supported by the fact that we can clearly have different opinions about what the best answer is without having to fall back into a position of the hopeless relativity of 'your answer is as good as mine' (which is an idea which runs quite contrarily to our adversarial system). There are two major recent papers on this topic. The longest and most difficult is 'Objectivity and Truth: You'd Better Believe It' in *Philosophy and Public Affairs* (1996). A somewhat shortened and simplified version is to be found in S Guest (ed), *Positivism Today* (1996) and it is entitled 'Indeterminacy and Law'.

The basic idea, put first in philosophical terms, is that a theory of truth in a particular domain of thought, such as law, or morality (or art), is a first-order, as opposed to a second-order, theory. That is not as difficult an idea as it might sound. It means simply that you don't have to suppose anything other than a straightforward argument of law to see what counts as law, as a *true proposition* of law. You don't have to jump to a higher plane – a 'second-order' – to 'look down' at ordinary legal argument to see whether it is producing a true conclusion. You don't, for example, say that because you can't prove many legal arguments to be right, they *can't* be right. All you have to do (I use this phrase advisedly) is to show again what these legal arguments are. Another way of putting it is to say that, to attempt to persuade a judge of the truth, cogency, etc, of your arguments there is no need to

persuade a judge of the truth, cogency, etc, of your arguments there is no need to do anything other than put those arguments to him.

A second-order justification, however, would say something like the following: the fact that there is no independent standpoint from which the objectivity and certainty of the law, or morality, can be assessed as true or false, means there can be no such thing as a right or wrong answer. Dworkin calls this imagined independent standpoint 'the Archimedean point' after the Greek philosopher who said he could lift the world provided he had an independent and sufficiently distanced fulcrum point. Dworkin points out, as he has so many times in the past, that the required Archimedean point doesn't exist to provide objectivity *to itself*, but he is much more concerned to say that, really, it is a waste of time to think that arguments about the 'objectivity' add anything to the actual arguments themselves. So, in 'Indeterminacy and Law' he says:

> '... when lawyers disagree, and there is no knock-down argument available to reconcile them, it follows that the case for neither side is better than the case for the other. There are an unlimited number of reasons why some but not all lawyers might think that one side had the better of a particular legal argument. Someone defending the view that no such reason can in fact tip the balance either way in any controversial case faces an enormously difficult task, much more difficult than that faced by someone who wants to argue for one decision rather than another in a particular case. How can he avoid appealing to some very general and abstract theory, like legal positivism? Someone defending a ... claim ... that there is never a right answer to any question about what we ought to do or how we ought to live ... has an even greater problem ... These are truly heroic claims, of vast theoretic pretension, and trying to dress them in the modest clothes of common sense or raw intuition is more comic than persuasive.'

Legal Concepts

26

The Analysis of Rights

26.1 Introduction: the place of law

Rights claimed in modern society have a contradictory quality about them. We can easily place strongly affirmed rights in direct conflict. For example, people claim the right to life yet there are others who claim a right to abortion; people claim the right not to be killed by another, yet there are also claims to a right to die; and people claim the right to free information, yet there is also a claim to privacy.

These are but a few examples. The claim to right is thus ultimately a claim to self-determination, which can produce logical contradictions and is itself in contradiction to the aspect of social control by law. However, the contradiction is one of degree. Thus, the issue of rights in the social context is one of balancing conflicting claims and determining which claims have priority.

The law has a special function within this framework. Law presupposes free choice at least to the extent that by implying that you ought to obey, you might otherwise choose to do something else. However, law restricts the way in which you may act in certain circumstances, even to the extent of physically restraining you. Thus, the law itself claims that the citizen ought to do as the law chooses, regardless of whether there are other non-legal reasons for doing otherwise. The origin of the right or authority to make law has been intermittently discussed, since this is the question of the authority to make law and of the obligation to obey it. However, it begs the question, does law extinguish individuals' claims to rights? The question may be broken down further:

1. Are there strong normative reasons for law to prefer an individual's choice to the prescriptions of legal norms? This must be considered the normative jurisprudential question of rights.

2. When a legal system concedes the existence of rights, what does this mean and what does a legal right do? This is a question of analytical jurisprudence.

26.2 Some contrasting views on rights

The legislators' duty, in ethical terms, is to society as a whole. Yet society is made up of interest groups and individuals. The immediate need of society may be seen as being in irreconcilable conflict with that of the individual. When should an individual or class claim of right be upheld in spite of the interests of the whole of society?

Marxism

The orthodox Marxist perspective on morality and rights stems from two premises:

1. Morality is an ideology that derives from the particular stage of development of productive forces of a society. Thus, Marxism cannot criticise the infringement of rights of workers in capitalist societies in moral terms. The critique of capitalism is a scientific one.
2. Man in socialist society requires no such ideology because he will naturally orient himself to social usefulness.

As a result, Marx views morality as relative to the particular stage of societal development and human rights as an ideology that alienates one man from another. As we have observed, rights presuppose restraint and conflict, mediating between them. Such alienation and mediation are seen by Marx and Engels as delaying revolutionary change to a society where conflict no longer exists. To adhere to a concept of rights, is to adhere to a maintenance of the status quo and unequal distribution. Marx denies therefore that there are strong normative reasons for rights that can be accepted by law, since law merely shields the interests of the dominant class.

Jeremy Bentham

Bentham, as has been observed in the Chapters on imperative theory and utilitarianism, completely rejects the concept of rights as anything other than fantasies of the mind. To Bentham rights derive entirely from the law and are legal constructs. However, it must be remembered that Bentham is sceptical about the concept of morality as a whole. Human beings act on the principles of pleasure and pain. It would seem to be vastly illogical that the interests of the rest of society should be subverted for the pleasure of an individual or class of individuals. The crude utilitarian perspective sees the legislators' duty as being solely the maximisation of pleasure in society, potentially at the expense of the rights of the minority.

However, this does not mean that all utilitarians follow this rather simplistic view. It is possible to demonstrate that a presupposition of weak rights is compatible

with the utilitarian perspective. We shall investigate the rights versus utility debate a little later on.

Bentham himself subscribes to the view that an individual should be granted the maximum independence that is conducive to the good of his fellows in society, but he still reserves the right of the state to intervene on behalf of the collective good.

Natural law

In Chapter 11 we discussed the views of natural lawyers, who tend to view natural law duties as ones that transcend legal duties. By appeal to natural law, we might have rights that exist independently of law that we would expect law to fulfil. However, it has been observed that natural law proofs tend to be open to empirical attack. In order to assert the existence of natural rights one needs to believe in natural law. Faith, either secular or religious, is a strong normative reason for an individual to expect rights, but not for an agnostic society to accept those rights. A further problem with natural law is that, historically, the distribution of rights has been uneven and thus a recipe for the denial of rights to some in favour of the privilege of others.

The facts of the 'human condition'

Both Hume and Hart suggest that there are certain empirical facts about the nature of the human condition that one would normally expect to see responded to in legal or moral systems. These amount to their respective theories of natural law. Inevitably, if we look at history, we can see that certain legalised actions have been ultimately detrimental to societal interests. Thus, genocide, torture and certain other extremes of state action in the name of society have served no particular benefit to society. In these terms, it is common sense for a historically informed legislator to avoid such excesses. Moreover, this is linked to a weak moral argument that law, while necessary to mediate between conflicting wills, should leave a certain amount of moral autonomy to the subject.

The rather common sense approach does, however, rely on practical reasoning and experience. This means that pragmatism can trump this conception of rights. The torture of a terrorist may save the lives of hundreds of potential victims of a bomb that he has planted. The argument against doing so is largely a moral one.

We shall explore interest theories, the chief modern theorist being Neil MacCormick, and Hart's will theory in greater depth later on, since they are concerned chiefly with the nature of rights actually found in law, rather than the reason why law should have a concept of rights.

It must be remembered from Chapters 20 to 23, however, that normative theories of rights, those of Rawls, Nozick and Dworkin, are to be found as part of their theories of justice.

26.3 Hohfeld's scheme of rights

Within the area of the analytical jurisprudence of rights the starting point for any study must be, according to Lloyd and Freeman, the work of Wesley N Hohfeld. Hohfeld's writing on the subject of rights was undertaken in the early years of the twentieth century and it could indeed be said with some justification that he has made a considerable though hardly acknowledged contribution to our understanding of law. In his work *Fundamental Legal Conceptions as Applied in Judicial Reasoning*, Hohfeld stated that the aim of his theory was to clarify different kinds of legal relations and the different uses to which certain words that are employed in legal reasoning are made. He sought to expose the ambiguities and to eliminate the confusion that surrounds these words. He was concerned to give meaning to the phrase 'X has a right to R' and to explain the set of jural relations that such a statement gives rise to. That objective can be achieved by the concept of right (which he also referred to as a claim); of privilege (liberty); of power and of immunity. These he saw as the lowest common denominators in which legal problems about rights could be stated. That proposition is one that is not without criticism. Indeed the contention of his critics is that while his scheme works for some propositions in which the phrase 'X has a right to R' could be fitted, it does not always work because his scheme could not take account of paternalistic criminal law. It is proposed to deal with this criticism in more depth below.

For Hohfeld these words (claim; privilege; power; and immunity) are to be explained in terms of correlatives and opposites, as each of these concepts has both a jural opposite and a jural correlative. These contain eight fundamental conceptions and all legal problems could be stated in their terms. They thus represented a sort of lowest common denominator in terms of which legal problems could be stated. This he did by method of the following:

Jural opposites – right/no right; privilege/duty; power/disability; immunity/liability.
Jural correlatives – right/duty; privilege/no right; power/liability; immunity/disability.

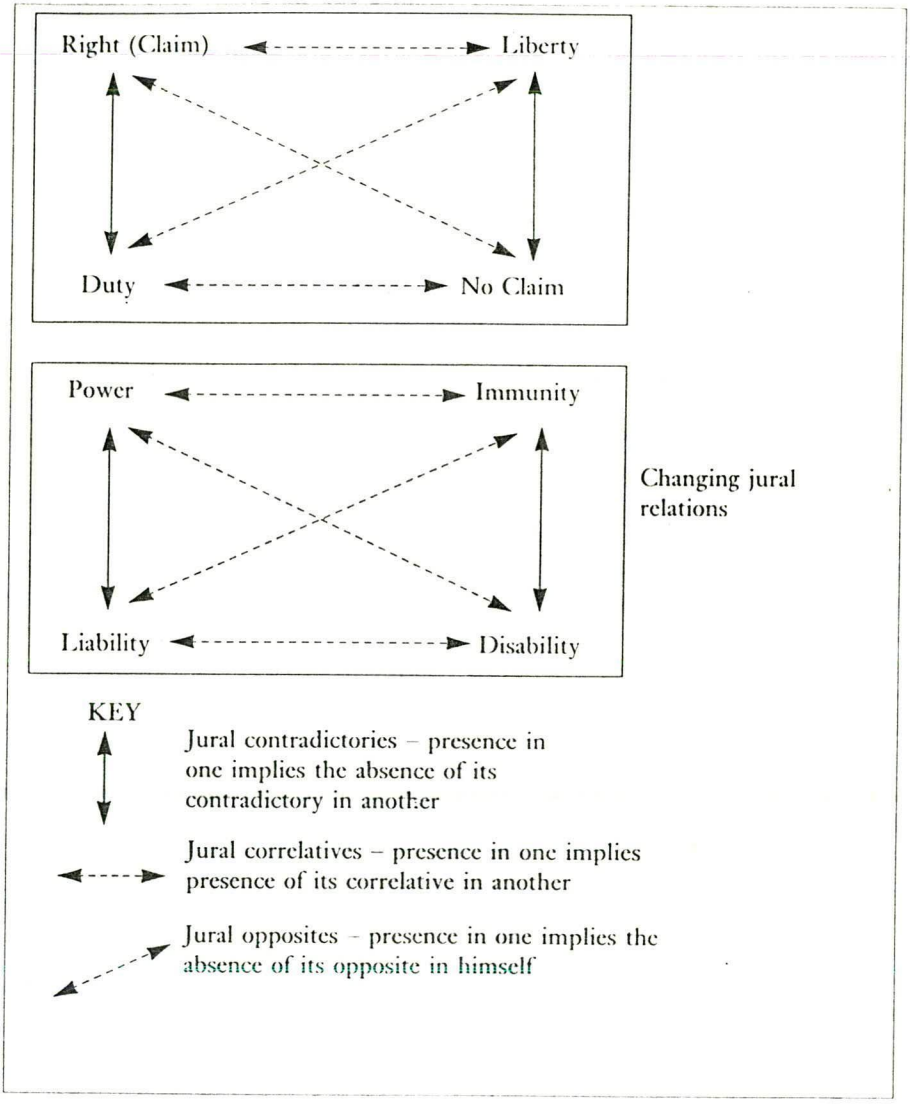
These terms can be defined as follows:

1. By a right (claim) he meant that everyone is under a duty to allow X to do R and that X would have a claim against anyone from everyone to enforce that right.
2. By a privilege/liberty he meant that X is free to do or refrain from doing that which is the subject of R. Y has no claim against X if X either exercises or refrains from exercising that liberty.
3. By a power he meant that X is free to do an act whether or not he has a claim or a privilege and that this act would have the effect of altering the legal rights and duties of others.
4. By an immunity he meant that X is not subject to anyone's power to change his legal position.
5. By a duty he meant that Y must respect X's right.
6. By no claim he meant that where X has a liberty Y has no claim that X should not exercise that liberty.

7. By disability he meant that the party has an inability to change another person's legal position.

It is important to emphasise that Hohfeld was examining legal rights and that the meanings attributed to his terms are technical and do not necessarily accord with their common usage.

Dias utilises a model developed by Glanville Williams which can be set out as follows:



The problem with this diagram is that while it appears to work on its face it does not use the same terminology as Hohfeld himself used and has therefore perhaps added to the confusion. On examination of the Glanville Williams diagram the arrows for jural correlatives are inverted with the arrows for jural contradictories. On the definitions offered by Hohfeld for his own terminology it would appear that a claim could not be regarded as the correlative of a liberty in the sense that the presence of a claim in one implies the presence of a liberty in another. It is submitted that the relation between a claim and a liberty is better described in Hohfeldian terms as a jural contradictory in that the presence of a claim in one implies the absence of its contradictory (a liberty) in another. This point is not however settled.

The aim of Hohfeld was to provide a model for the correct solution of legal problems and to make that solution easier and more certain. He urged that the judge and the legal theorist employ the above scheme in order to ensure greater understanding of these legal concepts. He wrote of the need to use the term right in a very strict sense and not indiscriminately to cover a privilege, power and immunity. Nonetheless, it would not be necessary to Hohfeld that the legal practitioner actually employ the terms claim, power, etc, so long as he thinks in terms of Hohfeld's scheme. It is thus possible to think Hohfeld without talking Hohfeld. This adequately deals with the criticism of Hohfeld that he has adopted an unusual terminology which it would be naive to expect the legal profession to adopt overnight. Indeed Hohfeld's scheme was developed seventy years ago and still nothing much has happened by way of the legal profession adopting his terminology in the effort to clarify legal problems. Nonetheless, his contribution has been quite substantial, although underrated to date. Hohfeld's scheme does, too, provide an excellent starting point for any theoretical discussion of rights.

Dias suggests that it is useful to view the relationship between the jural relations at rest and the changing jural relations in a *temporal* perspective. He argues that a change in the power/liability relation will have a knock-on effect on the claim/duty relation. The power/liability relationship would be anterior to the claim/duty relationship since the claim/duty relationship would be created or amended by the power/liability relationship and would reflect any change therein. An example would be in the adverse possession of land. Here a change in power/liability where formerly the tenant could exclude the squatter has now changed and the squatter gains title to the exclusion of the tenant. Thus T previously had power, that is, the facility to alter another's (S) legal status. On the running of time S now acquires power and T is under a liability. This would then lead to a change in the claim and duty relation.

26.4 Evaluation of Hohfeld's scheme

While Hohfeld's scheme of jural relations is useful not only for illustrating the different forms which the word right can take, it also illustrates the inter-

relationships between these words. It is thus useful for distinguishing between claims, liberties, powers and immunities but it is argued that it would also be both necessary and desirable to retain a general concept of right to denote institutions such as ownership or possession. As Cook, who was the editor of Hohfeld's work and generally sympathetic to his task, observes, Hohfeld mistakenly considers all rights as sets of any number of his four elementary rights, namely: claim, privilege, power and immunity. Rights are not sets of these. Their possession entails the possession of other rights or of powers and duties. For example, the concept of ownership includes rights of possession, transfer, sale, hire, use and enjoyment. Thus ownership creates a set of claims and powers. The concept of ownership can be seen as a set of rights. It does not denote the relationship between the owner and the tangible object.

His contribution has been useful although the difficulty is that it is not as widely used as he would have advocated. Nonetheless, as Lloyd and Freeman observe, it is the point to which all lawyers return. They perceive the value of his analysis in enabling the reduction of any legal transaction to relative simplicity and precision and in the enabling of the recognition of its universality.

Harris identifies three important advantages to his approach. First, that it enables real normative choices to be disentangled from verbal confusions. Secondly, that if lawyers and judges were to employ his terminology that was not too far removed from that already employed, then clarity would reign. The third advantage lies in their use. Hohfeld believed that juristic problems concerning the nature of compound concepts could be dissolved.

Although he has been criticised for insisting on correlativity in situations where correlativity is hardly present as, for example, in the criminal law, the implicit answer that Harris finds in his defence of Hohfeld is that all litigation cases involve two opposing parties and as such viewing these concepts as correlatives is in that frame quite meaningful. It does however make an explanation of rights *in rem* impossible. Nonetheless, there are important criticisms of Hohfeld's scheme. In that he purports to analyse fundamental legal concepts he does so without taking account of any concept of law. He fails to provide an explanation of the process by which those conceptions are given their legal character. He further assumes that there is only one concept of duty. It is said that this is because his examples are drawn from civil and private law.

In criminal law his scheme hardly works. This, it is submitted, is because of the nature of the duty under criminal law. While, as Harris observes in defence of Hohfeld, Hohfeld was concerned with the lowest common denominator in litigation, this would not in my view be applicable in a prosecution. The duty is not owed to the prosecution but to the society as a whole. That duty does not give rise to a right in anyone. Hohfeld's scheme is designed to cover one to one relations and not the relations between an individual and the society. Furthermore, it is suggested that with respect to paternalistic criminal law such as the laws that govern the wearing of seat belts in cars and the law of murder which forbids the defence of consent of the

victim, the nature of the duty is one that the individual owes both to the society and to himself. As such, when an individual has both a claim and a duty with regard to the same thing, the scheme would be without application.

Roscoe Pound in *Legal Rights* noted that some of Hohfeld's conceptions are without what he called juridical significance yet in a generally appreciative work Pound suggested that had Hohfeld lived he would have dealt with this point.

In spite of these criticisms, viewed in a chronological frame his contribution has been substantial. However, since the publication of his work there have been further developments and elucidations such as the works of Hart and MacCormick on rights. They benefited from having available to them Hohfeld's analysis.

26.5 The choice theory versus the interest theory

In this context the debate between Hart and MacCormick over the role and nature of legal rights is particularly informative. The essence of the debate should be viewed within its political perspective.

Hart's will theory

Hart views rights as legally protected choices. He emphasises the power or option of one person to waive someone elses duty. Thus having a right is to do with the legal or moral recognition of some individual's choice as being pre-eminent over the will of others as to a given subject matter in a given relationship. This is applicable in the civil law area in matters such as contract. The essence of the holding of a right is that the holder has the choice whether to waive the duty owed to him. The connection with Hohfeld's scheme of jural relations is apparent in that such a view assumes a correlativity of rights and duties. In this theory the choice could be expressed in Hohfeldian terms as the choice of whether or not to exercise that right or power or privilege or immunity. A problem with this approach is that it makes the enforcement of a duty conditional on the exercise of a choice or will of a person other than the person who is under the duty. Y will only be under a duty if X who has a right in respect of that duty decides to exercise that right. A difficulty that Hart readily admits with this approach is that it fails to take account of the fundamental rights of the individual as against the legislature. For this right Hart invokes the immunity as defined by Hohfeld.

MacCormick's interest theory

MacCormick criticises Hart's theory on the grounds that there are some rights which do not seem to involve the exercise of a choice at all. He argues that, particularly in the area of paternalistic criminal law, the law limits the power of waiver without destroying a substantive right. An example would be in respect of assault or of murder. The law will not admit the consent of the victim in defence to

a prosecution. MacCormick argues that if one cannot consent to assault it follows that one is not exercising a choice on the right to freedom of the person. MacCormick maintains that the nature of rights can be viewed as protecting the interests of the right holder.

Looking at the difficult example of the rights of a child MacCormick draws a distinction between the substantive right and the right to enforce the substantive right. He shows that the child possesses the substantive right to have its interest protected but lacks the right to enforce that right – the right to enforce is exercisable by the child's guardian on behalf of the child. Further the child cannot in fact or in morals or in law relieve his or her parents of their duty towards it. MacCormick then prefers the view of rights as protecting certain interests in the sense that either moral or legal normative constraints are imposed on the acts and activities of other people with respect to the objects of one's interest.

Hart admits that if rights are all about choice then a young child would not possess any rights in that sense. As to the question of the protection of the child, Hart maintains that rights are not the only moral basis for protection and that other factors such as humanity, love and compassion also provide the basis for protection. If that is so then there would be no need for a formal assignment of rights to the child on its attaining the age of choice. Until that assignment of rights the parent would act as the child could have acted had it possessed the power to choose.

Hart rejects the view that rights are legally protected interests because he maintains that the interest analysis does not explain rights independently of duties. If a right is merely a protected interest then rights can always be expressed as a reflex of duties. MacCormick gives an example of the right of succession in intestacy. He shows that such a right cannot be rephrased in terms of the rights of the personal representatives because the right vests at death, prior to these duties. MacCormick maintains that the idea of correlativity obscures the fact that duties are imposed in order to protect rights.

Nigel Simmonds in 'The Analytical Foundations of Justice', *Cambridge Law Journal* (1995) sympathetically (but also critically) discusses Hillel Steiner's book *An Essay on Rights* (1994). Particularly useful to Jurisprudence students in Simmonds' review is his admirably clear discussion of Steiner's equally useful discussion of the distinction between the choice and interest theories. Steiner's basic idea is that human rights are fundamental to all theories of justice, which is a view that has been going out of fashion in recent years with the growth of communitarian theories of justice in which communities, or groups, or families are the 'building blocks' of a just and good society. The attraction of the communitarian view is that, whereas rights-centred type theories do not, at first sight anyway, easily accommodate the idea of individual duties to the community (albeit to other individuals to respect mutual rights), communitarian type theories do. Raz, for example, thinks that people do not have rights to certain public goods such as the public culture displayed in art galleries etc (or, more mundanely, to the air) but nevertheless have *interests* in them which the community has a collective duty to maintain. Steiner, instead, claims:

> 'Rights are the items which are created and parcelled out by the justice principle. We learn something about justice by examining the formal or characteristic features of rights.'

Simmonds does not agree fully with this view since there are logical problems. He believes in the reductionism involved being able to re-describe all propositions in terms of rights accruing to individuals. Examples such as Raz's proposition that individuals do not have rights to public goods is a compelling one; nevertheless there is also difficulty in understanding what the purpose of public goods is (why they are 'goods') unless they are good *for someone*. The fact that a very large number of people benefit from public goods is a reason for denying them wholesale direct interest, eg in the form of granting them a formal right, and is an explanation of why we think that no one has a right to a public good. But it would not follow from this that there was no right of a more abstract kind. In a very basic sense, it is true that one has a right to be respected in decisions involving participation in public goods without having the right to a portion of a public right to be delivered. Simmonds points approvingly to Steiner's revival of the analytical method employed by Hart and the linguistic school of philosophy of the 1950s (and implicit in English analytical jurisprudence since the time of Austin). A close examination of the 'meaning' of 'right', Steiner thinks, will lead to insights about the nature of justice in general, in contradistinction to the avowedly evaluative approaches to enquiry about law engaged in by Finnis and Dworkin.

There may, however, be a problem with identifying the beneficiaries of a duty. In his book *Central Issues in Jurisprudence*, Simmonds uses the example of the crash helmet law whereby all people riding on a motor bike are under a legal duty to wear a crash helmet. Who is the beneficiary? Surely not the manufacturers of crash helmets? MacCormick may not be entirely correct in his contention that the power to waive a right is not a necessary part of a right but is just something that a right often includes. In support of that contention he demonstrates that in certain circumstances it is necessary to override freedoms – for example, in contract the freedom to contract the terms is overridden by the recent consumer protection legislation. Simmonds sums it up thus:

> 'Even if MacCormick has provided a convincing case against the correlativity of rights and duties, it is by no means clear that he has provided a convincing alternative.'

MacCormick does admit the importance of the will theory in the explanation of rights. He put it thus:

> '... it cannot be denied that the central point of the theory is that apart from children and incapacitated persons the holder of a legal right is empowered in law to choose whether he should avail himself of his right on a specific occasion by insisting on performance of the correlative duty'.

If that is so it might be assumed (albeit wrongly, it is submitted) that Hart's and MacCormick's theories are compatible, but that would be to fall into the linguistic trap which was so much the concern of Hohfeld. For MacCormick the difficulty is

in the absence of choice with regard to children's rights, the argument going that those rights are among those referred to in a footnote by Hart in his notion of immunity rights dependent upon individual benefit. In the Hohfeldian sense the rights of children as envisaged by MacCormick are claim rights whereas Hart's are immunity rights. Hence while both rights are fundamental and important they have different lowest common denominators.

27

Legal Personality

27.1 Introduction

A right is not the only legal concept to have attracted much jurisprudential discussion. Another such concept is that of legal and especially corporate personality. Why are certain bodies treated in law as persons and some bodies (trade unions, partnerships, unincorporated associations) generally not? What, if anything, does it mean to say that a company is a person?

27.2 Different types of legal personality

There are three types of personality recognised in English law; in relation to any question, you should ask yourself if it is about one or all three.

Human beings

No distinction is drawn in law between legal and natural persons. Hohfeld sees human beings as merely a multitude of claims, liberties, powers, etc. But it should be noted that the notion of a human being is more flexible than might be thought. We shall examine some of these:

A foetus

What is the legal status of an unborn child (re pre-natal deformity? child destruction?) The example of the legal personality of a foetus has raised interesting and emotive questions recently. In the case of *C* v *S* (1987) one such question arose

for consideration. Briefly, in that case a man who claimed to be the father of a foetus attempted to prevent the mother of the foetus from proceeding with an abortion after their relationship broke down. His grounds were to invoke the criminal law against the destruction of a child capable of being born alive (s1 Infant Life Preservation Act 1929). The rather controversial interpretation given to that phrase by the House of Lords need not detain this text. What is of importance is the observation that it was the father (as an interested person) who brought the action and not the foetus, yet if the foetus had been deemed to be a legal person it could have brought the action itself. Practical problems of instructing solicitors, etc, from the womb can in this legal system be overcome – there are procedures to enable the incompetent to be party to actions. The implication though is wide. The foetus would have a separate legal personality from the mother carrying it. The mother would merely be a walking incubator for another legal person. The mother would owe that person a duty of care that would give rise to that person having a cause of action where, for example through smoking cigarettes, the mother caused the foetus damage. If the foetus were a legal person then it would be party to an action to prevent an abortion, etc. I would imagine that other factors will be considered in seminars – enough for present purposes to raise the questions. By way of anecdote for those interested, I understand that although the House of Lords held that the foetus could be aborted as at the stage of gestation it had reached, it was not capable of sustaining life if born, the mother continued with the pregnancy and the child is being brought up by the father. (A happy ending?)

A dead person

Legal personality extends to those humans who are alive and of an existence independent of their mother. What is the position with regard to the dead? They have legal interests, such as that their wishes as expressed in their wills are carried out, for example. The law has studiously avoided any definition of death – see *R* v *Malcherek and Steel* (1981) – probably for the very sound reason that advances in medical science and technology would outstrip the capacity of the law to keep pace and we would arrive at a situation, as we have for example in criminal law, where the definition of insanity became fossilised in 1843 (*McNaghten's Case*) in spite of very considerable advances since then. So what constitutes legal death is not clear. I think the proposition stated at the start of this sub-section to the effect that so far as natural persons are concerned a prerequisite of legal personality is independent live existence is true. The dead may have certain rights, such as to have their property disposed of according to their legal wishes, but that is as far as it goes.

A married couple

The example that used to be used was that a husband and wife were treated as one person for certain tax matters, for example, mortgage interest relief and the filling of tax return forms that required a wife to disclose to her husband all the sources of

her taxable income, because the husband was under the legal duty to declare that income of his wife to the Inland Revenue. That position has been changed in the Finance Act 1988 although aspects of it have not at the date of writing entered into force. I do not envisage this rather antiquated rule that regarded a wife as an appendage of her husband being revived.

Status

Another flexible aspect is that there are different relationships to think about – status (parent, slave, consumer) and capacity (the same person can have two or more in some factual situation; trustee and beneficiary, shareholder and employee and company director, for instance).

Corporations sole

A corporation sole is a person with a perpetual existence, that is, an office, the personification of an official capacity. Examples are parsons, bishops, the Crown (the Queen has a different personality for each country where she is the monarch). The main rationale behind the corporation sole is that the continuity of jural relations, such as the holding of property, is made possible. This need hardly detain us further.

Corporations aggregate

These are companies or other corporations created by charter, statute or under the Companies Acts. They are treated as persons in law unless the contrary is stated (statutes use individuals if they mean humans and unincorporated associations but not corporations). Some unincorporated associations are given some of the incidental benefits of corporations but they are still not persons. Partnerships, for example, can issue writs in their own name and can make contracts, but the individual partners remain fully liable as individuals.

27.3 Is legal personality a useful concept?

The flexibility of treatment given to the notion of a human being is useful and corporations sole have their limited effectiveness allowing in the continuation of property ownership and contractual relations. This question as to the usefulness of the concept of legal personality is, though, most relevantly considered with regard to the corporations aggregate.

The uses of corporations aggregate

These have been stated as:

Convenience
The convenience offered by conferring powers and liabilities on a unit rather than on each individual shareholder involved (imagine suing British Telecom if it were otherwise!).

Limited liability
Shareholders do not attract liability except to the extent of their respective shareholdings, and directors and employees are only liable for their personal negligence; this is subject to the frequent requirement of personal guarantees, from participants in small companies.

Perpetuity of succession
This applies on death, retirement, sale of shares.

Ability to sue
Can sue or be sued, can own property (this is really an aspect of Chapter 26).

Separate ownership
Ownership and control can be separated, allowing investors to risk their money but under the control of expert management. In many public quoted companies, ownership and control are totally divorced in this way.

Other advantages
Generally, an individual trader can escape personal liability (subject to personal guarantees) and it is easier for a company to raise capital than for a sole trader. Note though that the courts do sometimes lift the corporate veil (there is a list of instances of this in Dias). Note also that some of the advantages can be achieved without a separate personality being used. For example, writs can be served on partnerships, and property is often held by only some of the partners. This allows ownership to be passed more easily. Also, a big partnership will often separate its management from the bulk of the owning partners. Think also of the special treatment of trades unions and employers associations in English law.

Problems with corporations aggregate

Corporations aggregate also raise the following problems:

Groups
English law has difficulties in dealing with the idea of a group of companies. For most purposes, they are treated as separate units rather than as a collective entity, which is most unrealistic. Some inroads have however been made into this problem; group companies now submit consolidated accounts and are taxed as a unit, for example.

Inflexibility

Even small companies have to fulfil statutory requirements suitable for much larger outfits. The present government is committed to relaxing some of these requirements and it is already the case that small companies have to submit less complete accounts, for instance.

Unfairness

Small creditors never have security and so can lose heavily in an insolvency. There are some restrictions in the Insolvency Act 1986 on directors involved just setting up another, similarly named company in such cases (only time will tell if they are effective) and there are various possibilities of penalties or civil remedies (including disqualification) against directors involved in an insolvency – see the Companies Directors Disqualification Act 1986. But these generally will not benefit small unsecured creditors, who can't afford to rely on them. It is a myth that a customer or supplier is safer dealing with a limited company than an individual trader or partnership – often the reverse is the case.

Inconsistency

It is not at all clear why some legal rules and regulations apply to all companies but not to other (often larger) organisations which organise themselves as partnerships. Often the choice of business medium is based on taxation considerations rather than on which medium is more suitable in terms of its inherent characterstics.

27.4 The theories: what theories are used to explain legal personality?

There are four main theories for us to consider, albeit in each case rather briefly. We will try to analyse for each theory which elements of the law it can, and cannot, account for.

Fiction theory

First, we will look at the fiction theory, supported by von Savigny and in England by Coke, Blackstone and especially Salmond. Juristic or artificial persons are only treated as if they are persons, under this view. They are fictitious, not known as persons apart from the law. The law gives them proprietary rights, grants them legal powers and so on, but they have no personality and no will (except to the extent a will is implied by the law). This is an obviously flexible viewpoint, since it can account for any apparent inconsistency in legal treatment by simply saying that they are only treated as persons to that extent. The doctrine of *ultra vires*, under which a company cannot do anything not authorised by its memorandum of association might be thought to support the fiction theory, on the basis that the law only gives personality to the limit of the memorandum, and so might the doctrine that a

company is separate from its members, epitomised in the leading case of *Salomon*. This case shows that the law treats the company as a separate unit, even though in fact it is not, especially in the one-man company cases like *Salomon*.

Further support for this theory could be claimed from the criminal law, which originally accepted that a company could not commit a criminal offence which depends on mental intention. The fiction view explains this on the basis of the will of the person only being that given by law, and therefore presumably being limited to lawful intention. Recent developments show a more pragmatic and sensible approach to the question of corporate liability, with companies being subject to more criminal liability (and also subject to liability for the torts of their servants). Also, the cases where the law allows the corporate veil to be lifted aside can be explained as limitations on the grant of the fictitious personality.

Acceptable explanations, then, are provided by the fiction theory for many aspects of company law (although many of them can be explained acceptably by other theories, see below).

However, no explanation is given of why the law uses the idea of personality; is there an essential similarity to real persons or not? Hart has emphasised some of the faults identified in relation to this theory, particularly the illogicality involved in denying that a company can commit certain crimes because it has no mind.

Some other theories are similar to and bound up with the fiction theory, notably the concession theory (that legal personality flows from the state) and the symbolist theory of Ihering.

Hohfeld's theory

This theory is not mirrored in English writing on the subject. Since only human beings have juristic relations, one must, according to Hohfeld, explain companies in a complex way by looking at the capacities, rights, powers and liabilities of the individuals involved. This view is clearly related to Hohfeld's analysis of rights. However, it again fails to give us an explanation of why the notion of a company is used, the notion of a separate personality.

Realist

This view sees an artificial person as a real personality, having a real mind, will and power of action. It is associated with Gierke, Dicey, Pollock and (though Hart doubts it) Maitland.

If independent power of action was the only requirement of our definition of a person and personality, perhaps an artificial person would qualify (but has a company really got a power of action independent of its members and officials?); surely though there is something more. To say a corporation is a real person implies an individuality, and that implies some consciousness, experience, inner unity. Some

groups may seem to have such a unity and consciousness. Perhaps one could talk of such a feeling over the reaction to the Somalian famine crisis, for instance, a corporation sole (consisting of successive holders of one office) hasn't a consciousness, nor has a multi-national company, nor even a small company? Perhaps a university might be thought to fit?

In any case, even if the legal personalities could be counted as real persons, a further problem arises. If a two-man company is a person in reality, why not a two-man partnership? If a one-man company, why not a one-man business? If a university, why not a private law college, one that is unincorporated? The realist theory fails to explain why the legal definition of personality does not match the extended realist definition.

Returning to some of the aspects of English law already considered, realist theory can account for the *ultra vires* doctrine (the real personality constituted by the company as set up by its documents), albeit rather weakly (isn't it a weakness to have to refer to legal documents to establish the limits of reality?); but it can't successfully accommodate the tearing aside of the corporate veil. If the company is a real entity distinct from its members, surely it should always be viewed as such and not sometimes viewed as a collection of its members?

Finally, realism can account for those instances where criminal law applies to a company: can it account for those when it doesn't (if a board meeting orders an execution, the company isn't guilty of murder: why not)? The reason why it would not be guilty of murder in the likelihood of a board resolution so ordering is that it is incapable of forming the necessary *mens rea* for murder. Obviously, considerations as to suitable penalty will also be relevant (it is impractical to imprison a company!). The recent suggestions of possible prosecution whether public or private against a ferry operating company for corporate manslaughter, as a result of suggestions in the inquiry into the events at Zeebrugge, demonstrate an actual example of the criminal law responsibility of a corporation.

Hart has raised some additional points. The theory (as with the fiction theory) has illogical barriers; for example, it has been suggested that a company cannot be bound by an agreement with another company because that would be degrading, and the realist view has difficulties with a one-man technical company – isn't such a company not a real entity but just a convenient device for the individual proprietor?

Linked to this theory is the organic theory, the name of which is quite accurate and suggests that a company acts through the various organs that constitute it. In this theory the board of directors will take decisions concerning the day to day running of the company whereas the general meeting will take decisions concerning the constitution of the company. So long as the proper decision is taken in the proper way by the proper body it will be considered at law as the act or decision of the company.

Purpose

The final view is the purpose theory developed by Brinz and, in England, by Barker. On this view, only human beings are persons, but the law protects certain purposes other than human beings. The creation of artificial persons just gives effect to a purpose (for example, a charitable corporation is created to give effect to various devices by which the law aids the charitable cause). So company property is held not by a person, but for a purpose.

This view has a fundamental flaw. It does not answer the question. It is obviously true that companies and other artificial legal persons are given their status for a purpose (or various purposes). The question remains, why call them persons? What aspect of these entities makes them so akin to real people that the law uses the same name and to a great extent applies the same rules?

A purpose view can explain the *ultra vires* doctrine (a company is limited to its express purposes, as mentioned in the memorandum), and even the tearing aside of the veil (the countering weight of other legal purposes), cannot explain the concept of an artificial person.

Our conclusion at this point is that none of the various explanations given of the nature of corporate personality is satisfactory.

27.5 Do the theories obscure?

Both Paton in *Jurisprudence* and Hart in his inaugural lecture at Oxford, *Definition and Theory in Jurisprudence*, think that the answer to the question is yes, because the theorists, in trying to ascertain what is the nature of corporate personality, are asking the wrong question. Paton writes that seeking the essence of, the connecting factor between, the various different types of legal persons, natural and artificial, is the wrong approach because there is no connecting factor except the similarity and treatment meted out by the law to the different persons.

Hart's views are somewhat different. The slim booklet containing his inaugural lecture (which is also included in his *Essays in Jurisprudence and Philosophy*) is well worth reading. Briefly he considers that the question 'What is corporate personality?' is seeking the wrong type of definition.

Just as with other concepts found in law, such as a right or a duty, corporate personality has no straightforward connection with the world of fact, nothing to which it corresponds. It should not, therefore, be defined in the same way as the concept of the table or chair you are sitting at or on, since these concepts do have an object to which they correspond in nature – they exist and can be touched in the real world of scientific cause and effect. Hart goes on to say that the theories which we have looked at above are often in the clouds and do not deal with practical realities.

What then is Hart's alternative? Well, he says, picking up an idea that he attributes to, among others, Bentham, corporate personality and the other legal concepts should be defined not by looking just at the words themselves but rather by considering a characteristic sentence in which they appear and then explaining the conditions under which the words are used – under what conditions does the law ascribe liabilities to corporations? By this method we can avoid questions such as those relating to a corporation's supposed will.

Hart in *Definition and Theory* imagines an innocent lawyer from Arcadia to whom the notion of a legal or corporate personality is introduced for the first time. He would learn what types of legal personality there were and the forms of statement in general use by which rights were ascribed to Smith & Co Ltd, in circumstances in some ways similar to and in some ways different from those in which they were ascribed to Smith as an individual. He would see that the analogy was sometimes thin but that given the circumstances set out in the Companies Acts and in the general law the statement Smith & Co Ltd owes White £10 applied as directly to the facts after its own fashion as Smith owes White £10.

On his return to Arcadia he would tell of the extension to corporate bodies of rules worked out for individuals and of the analogies followed and the adjustments of ordinary words involved. He would, in short, have explained corporate personality without any need to get into the confusing and obscure theories which I have set out above. In Hart's words we could make the simple Arcadian feel the theorist's agonies only by inducing him to ask what is Smith & Co Ltd and not to admit in answer a description of how, and under what conditions, the names of corporate bodies are used in practice, but instead to start the search for what it is that the name taken solely describes, for what it stands, for what it means.