It is a frequent complaint against academic moral philosophers that they inhabit an ivory tower where they meditate upon the analysis of ethical concepts or explore the logical semantics of the language of morals without ever descending to earth to apply their abstract studies in giving practical advice as to how to solve actual moral problems. The reply we academics usually make to this accusation is that our critics are confusing the function of an ethical philosopher with that of a moralist: the function of a moralist is to advise or to exhort, that of a philosopher to think about ethics, or, indeed, to think about thinking about ethics. But I, for one, am sufficiently sensitive to this criticism to be prepared to moralize about anything which my philosophizing has given me reason for moralizing about—especially when, as in the matter about which I shall venture to advise you today, the exposition of the reasons for the advice I shall offer as moralist are exactly the considerations which will be involved when, for a few moments towards the end of the lecture, I shall re-ascend the ivory tower to philosophize upon the ethical concepts I have employed.

The concept with which I shall be principally concerned is one which is given various names in different contexts—*Justice*, sometimes, in one of its many senses; *equitable distribution* in the context of welfare economics; *Égalité* if one is manning the barricades in its defence. I shall be very British and will call it *fair*

Cambridge University Press
978-0-521-11351-9 - Theory of Games as a Tool for the Moral Philosopher
R. B. Braithwaite
Excerpt
More information

play or *fairness*. The situations in which it is involved are those in which we, as social beings, find ourselves every day of our lives. We cannot get what we want except by co-operating with other people whose wants are different from ours. It would be reasonable for Robinson Crusoe to aim simply at maximizing his own satisfaction while he was alone on his island. But when Man Friday arrived, co-operative action became possible, and in this Friday's wishes would have to be taken into account. Can the philosophical moralist give any advice to people with different aims as to how they may collaborate in common tasks so as to obtain maximum satisfaction compatible with fair distribution? I need not elaborate upon the practical importance of securing such collaboration in all spheres of life, domestic, social, national, international.

Moral philosophers have, for the most part, ignored this question. They have supposed that, unless people can agree at any rate upon the proximate ends which they all wish to pursue, no rational basis for common action is possible. Many philosophers have pointed out that, however wide a divergence there may be about ultimate ends, there is a large measure of agreement as to proximate ends (such as the maintenance of law and order, or the preservation of individual freedom) which are necessary conditions for many different ultimate ends, so that the matter is not quite as hopeless as it appears to be. But they have left any serious study of the question to the welfare economists who, heirs of the Utilitarian tradition, have evolved theories of production and distribution based upon the assumption that the ends desired by different people—their 'utilities'—can be compared with one another in terms of common units (frequently monetary units) which can be transferred from one person to another. Welfare economists, however, whether they regard themselves as describing the actual working of an economic system or as prescribing a more ideal functioning, have come in for a great deal of criticism recently from their fellow-economists based largely upon the impossibility of any inter-personal comparison of utilities. So many economists today are as negative in their attitude to our problem as are the philosophers—though with more excuse, for they at least have tried out one line of approach.

I propose today to try another line of attack, suggested to me by thinking about the new mathematical discipline called Theory of Games. Though this approach will presuppose that each person can compare his own preferences for alternative courses of action, it will not presuppose that the preference scale of one person can be compared with that of another; a method of inter-personal comparison will arise naturally in the course of the argument. My treatment will be ethically neutral between the collaborating parties; it will not suppose that one code of values is better than any other, nor will it require that any of the collaborators should attempt to convert his colleagues to his own system of valuation. The recommendations which I shall make for sharing fairly the proceeds of collaboration will therefore be *amoral* in the sense that they will not be based upon any first-order moral principles; but the recommendations themselves will constitute what may be called second-order moral principles giving criteria for *good sense*, *prudence* and *fairness*, principles in some way analogous to the 'supplementary Principle' for 'Just distribution of happiness' which Henry Sidgwick, my 'great-grandfather' in this Knightbridge Chair, found necessary to sweeten the pure milk of the Utilitarian gospel.[1]

After these preliminaries, let me attack the problem directly. But first I must limit my objective to situations in which there are only two persons concerned in the collaboration. Although this is a serious limitation, the two-person situation covers more than may appear at first sight, since 'person' may be taken in the extended legal sense in which a corporate body—a University, a Company, a State, any social group which decides and acts collectively—may be said to be a person. Why it is necessary to start with the two-person case is that, if a group of more than two are involved, the possibility arises of forming subgroups within the group—cabals, coalitions, trusts, cartels, etc.—with mutual assistance pacts designed to promote the interests of the subgroup at the expense of those outside it. This possibility introduces serious complications into the problem, how serious has been shown by the Games Theorists. So my limitation to two persons is essential to my treatment.

But my second limitation is made purely for expository convenience. My two persons—let us call them Luke and Matthew (L and M)—will be supposed each to have only two possible ways of acting in the collaboration situation: Luke can act either in one way or in one other; Matthew can act either in one way or in one other. This limitation is inessential, since the mathematics can be straightforwardly, though laboriously, extended to cover situations in which each collaborator has the same finite number n of possible ways of acting.[2] But to expound the $n \times n$ case would require covering several blackboards with sets of simultaneous equations, whereas in the 2×2 case simple geometrical arguments can take the place of heavy algebra. So I shall limit myself in this lecture to the case in which Luke has two possible ways of acting, L_1 and L_2, and Matthew has two possible ways, M_1 and M_2, so that the results of their combined actions may be represented by a four-square table in which Luke has the choice between the upper and the lower row, and Matthew the choice between the left-hand and the right-hand column, and where the four τ's represent the outcomes of the four combinations of these choices:

	M chooses between: M_1	M_2
L chooses between: L_1	τ_{11}	τ_{12}
L_2	τ_{21}	τ_{22}

But you will probably prefer to think about the *pure invariant casuistry of two-person* 2×2 *fair collaboration* in terms of its application to a concrete

Cambridge University Press
978-0-521-11351-9 - Theory of Games as a Tool for the Moral Philosopher
R. B. Braithwaite
Excerpt
More information

example, and I shall continue my discussion entirely in terms of this example. I shall try to make it as realistic as possible, but of course various simplifying assumptions will have to be introduced in order that the example may be paradigmatic of the collaboration situation in general.

Suppose that Luke and Matthew are both bachelors, and occupy flats in a house which has been converted into two flats by an architect who had ignored all considerations of acoustics. Suppose that Luke can hear everything louder than a conversation that takes place in Matthew's flat, and vice versa; but that sounds in the two flats do not penetrate outside the house. Suppose that it is legally impossible for either to prevent the other from making as much noise as he wishes, and economically or sociologically impossible for either to move elsewhere. Suppose further that each of them has only the hour from 9 to 10 in the evening for recreation, and that it is impossible for either to change to another time. Suppose that Luke's form of recreation is to play classical music on the piano for an hour at a time, and that Matthew's amusement is to improvise jazz on the trumpet for an hour at once. And suppose that whether or not either of them performs on one evening has no influence, one way or the other, upon the desires of either of them to perform on any other evening; so that each evening's happenings can be treated independently. Suppose that the satisfaction each derives from playing his instrument for the hour is affected, one way or the other, by whether or not the other is also playing; in radio language, there is 'interference' between them, positive or negative. Suppose that they put to me the problem: Can any plausible principle be devised stating how they should divide the proportion of days on which both of them play (τ_{11}), Luke alone plays (τ_{12}), Matthew alone plays (τ_{21}), neither play (τ_{22}), so as to obtain maximum production of satisfaction compatible with fair distribution?

The relevant data will be the extents of the preferences of each for the four alternatives. We will suppose that Luke's first preference is for him alone to play (τ_{12}), his second preference is for Matthew alone to play (τ_{21}) (since he quite likes hearing jazz when he is not playing himself), his third preference for neither to play (τ_{22}), while the thing he likes least is for Matthew to trumpet while he himself is playing the piano (τ_{11}). But, except in the obvious cases, a mere ranking of Luke's preferences in an order of preference is not a strong enough premiss. In order to be able to say much of interest, it is necessary to assign numerical values, not to Luke's preferences themselves, but to the ratios of his relative preferences for one outcome rather than for another. This can be done in the following manner. Suppose that Luke has no preference, one way or the other, between the two possibilities (*a*) of both of them always remaining silent, and (*b*) of Matthew alone playing on one-third of the evenings, and of their both playing on two-thirds of the evenings, on the average. Luke's indifference between τ_{22} and what amounts to a probability combination of τ_{21} with probability $1/3$ and τ_{11} with probability $2/3$, can be used to define what is meant by saying that the extent to which Luke prefers

τ_{21} to τ_{22} is twice the extent to which he prefers τ_{22} to τ_{11}. Suppose further than Luke is indifferent between the two possibilities (*a*) of both always remaining silent and (*c*) of himself alone playing on one-sixth of the evenings and of them both playing on five-sixths of the evenings, on the average. This will similarly determine that the extent to which Luke prefers τ_{12} to τ_{22} is five times the extent to which he prefers τ_{22} to τ_{11}. These two ratios will enable a scale of numbers to be assigned in a consistent way to Luke's valuations of the four alternatives and of all probability combinations of them, so that his preferences can be *measured* in the weak sense of measurement in which both the zero and the unit of measurement are arbitrarily chosen. This weak sense is that in which temperatures are measured by thermometers graduated according to, for example, the Centigrade scale. It implies that any result we get by using the scale of numbers measuring Luke's valuations must be unchanged if we select a different zero and a different unit for the scale, just as any scientific law about temperatures must be expressible indifferently in degrees Centigrade or in degrees Fahrenheit. In mathematical language, the only properties of the numbers that can be relevant are those which are invariant with respect to any linear transformation $x' = ax + b$ with a positive. When we represent, as it will be convenient to do, Luke's valuations by points along a straight line, what is determined by the weak measurement is merely the ratios of the distances between the points, so that everything we want to say about the points must be invariant for any uniform translation of the points along the line, and for any uniform expansion or contraction of the line, i.e. for any change in the zero point or in the unit of the scale.

This weak method of measuring preferences, by equating the number to be assigned to one valuation with the number to be assigned to a probability combination of two or more valuations, is implicit in the use made of the notion of *mathematical expectation* by writers on probability since the time of Leibniz; it appears explicitly in F. P. Ramsey's suggestions for a calculus of degrees of belief read to the Cambridge Moral Science Club in 1926 and published posthumously in 1931. Its appropriateness has been the subject of a great deal of discussion in recent economic literature since von Neumann and Morgenstern elaborated it in 1944 for use in Theory of Games, the discussion centring round the fact that the probability-combination-indifference method (as I shall call it) does not allow for the love of gambling, or of insurance, for its own sake. But whether or not the method is altogether suitable for measuring preferences with respect to possible alternatives on a unique, unrepeatable occasion, e.g. if Luke and Matthew are birds of passage and are touching down in the flats on one evening only, it is the obvious way to measure their preferences if any one evening may be regarded as one of a long series of evenings in each of which the relevant conditions are exactly the same. So I shall adopt the probability-combination-indifference method without further question; no other method will allow of the use of Theory-of-Games techniques.

This method of measuring preferences, it must be emphasized, does not presuppose anything like the adding together of separate units of happiness which has been such an excellent cudgel with which to beat Utilitarianism in its earlier forms. It will be convenient, and a proper tribute to the economists, to refer to the numbers, chosen in accordance with the probability-combination-indifference method to measure Luke's valuations, as L-*utilities*. But to use this language is not to attach any special significance to a zero L-utility or to a unit L-utility; any result obtained will be independent of how the unit and zero are assigned; it will depend only upon the ratios in which the L-utility differences stand to one another.

If we assign M-utilities in the same way to measure Matthew's valuations, we can represent (as in Diagram I, p. 13) the pairs of valuations of the four possible outcomes of the collaboration situation by four outcome points (the four *base points*) on a Cartesian plane, where for each outcome point the horizontal or x co-ordinate represents Luke's valuation and the vertical or y co-ordinate Matthew's valuation of the outcome represented by the outcome point. Luke will wish to secure that the outcome point jointly chosen should be as far to the right as possible, Matthew that it should be as high up as possible. Since the origin of the co-ordinates and the scales along both the horizontal and the vertical axes are arbitrary, the only geometrical properties that can be relevant are those which are unchanged if the figure is moved without rotation or if it is expanded or contracted along either or both the axes. To assist in appreciating