

## DISCUSSIONS

### COMMENTS ON JOHN C. HARSANYI'S PAPER

Act utilitarianism (AU) and rule utilitarianism (RU) are defined rather vaguely and ambiguously in most of the literature. I want to argue that in some interpretation of the terms they are indeed equivalent, in others they are not. Let us think of a game  $S$  played by  $n$  persons  $1, \dots, n$  who have a choice between actions  $f_1, \dots, f_m$ . On the set  $X$  of results  $n$  utility functions  $u_i$  are to be defined. Then AU cannot be sensibly defined by saying: "An action  $f$  is *right* (in  $S$ ) if it will maximize social utility (which for a result  $x$  is defined by  $U_s(x) = \sum u_i(x)$ )". For what action will maximize  $U_s$  depends on who is acting and on what the other players do. So we can only say: "On condition that  $p$  ( $p$  being a set of possible outcomes) it is right for  $i$  to do  $f$  iff that is compatible with achieving an optimal  $p$ -result (with respect to  $U_s$ )." By generalizing over persons  $i$  we obtain a concept of rightness for actions, and by choosing a tautologous  $p$  we get concepts of *prima facie* rightness for acts and actions.

If we define RU as maintaining

- a) An action  $f$  is *right* if, realized by everyone, it will maximize  $U_s$ ,

then clearly RU and AU (*prima facie* version for actions) are not equivalent and the cases of non-equivalence are also cases of inadequacy of RU, as the following example shows, where  $f_2$  is forbidden according to RU, but allowed according to AU.

1 \ 2	$f_1$	$f_2$
$f_1$	2	5
$f_2$	5	1

However, if we consider RU as an answer to the question, what general practice should be adopted in a society, and formulate its thesis as

- b) A *general practice* is right if (with respect to other general practices) it will maximize  $U_s$ ,

then it is equivalent with AU, since according to AU an action  $f$  is right for anyone, on condition that everybody does  $f_1$  or . . . or everybody does  $f_m$  iff the result where everybody does  $f$  is optimal. And in this sense the equivalence thesis seems to have been intended by D. Lyons in 'Forms and Limits of Utilitarianism' (1965), for instance.

Now Professor Harsanyi has formulated AU and RU as principles that apply to different situations. According to his definition either what all the others do is given – then RU does not apply – or it is not – then AU does not apply. But we can compare AU and RU as to their relative merits only if there are at least some situations to which both apply. If we understand RU as a decision rule under the assumption that all will use the same strategy then a comparable form of AU will have to use the same assumption and it is then equivalent with RU, as I have pointed out. Consider Harsanyi's Example 1. He says that RU will require all voters to vote, but AU will require each voter to vote only if he expects all others to vote. But this is only true according to my first version of RU, not according to his definition, on which RU only applies under the assumption that either all vote or all do not vote. (This is a questionable assumption in a non-cooperative situation like voting since it implies that whatever I do, by some hidden machinations the others will do, too.) Now, on this same assumption AU also tells me to vote for if I do so, according to the assumption, the others will vote, too.

*Universität Regensburg*

FRANZ V. KUTSCHERA

Manuscript submitted 29 June 1976

Final version received 1 April 1977