According to "individualistic" ethical doctrines, moral norms are applicable only to individual agents. No moral norms apply to groups of agents. Thus, according to what might be called Individual Egoism, each (individual) agent ought to maximize his utility and there is nothing that a group of agents ought (or ought not) to do.

Österberg (1988) pleads for a collectivistic approach to morality. (Unless otherwise indicated, all page-references below are to this book.) In particular, he argues that the most defensible form of egoism should not only tell each of us what to do but also tell us what we ought to do. He also claims that collective norms should take precedence over individual ones. An individual ought to do one's part in an action-pattern that is prescribed for the group -- provided that other members of the group do their part. Roughly, you should cooperate with others as long, but only as long, as they cooperate with you. You should be prepared to cooperate even if you would gain by defection. On the other hand, you should not let the others take advantage of you. There is no reason to be a sucker.

This combination of collectivistic ideas with an essentially self-centred moral view yields Collective Egoism -- the most defensible form of egoism, according to Österberg. However, he finally argues that ethical egoism, even in this most defensible form, must be rejected -- no acceptable ethical doctrine can prescribe "that people ought to treat themselves differently from others" (p. 196).

In the present paper, I shall abstain from discussing this "fundamental" issue. Instead, I am going to consider whether Collective Egoism achieves what it is supposed to achieve: whether it has the advantages that Österberg claims on its behalf. In particular, I am going to question his claim that Collective Egoism, unlike other forms of egoism, avoids violations of the principles which he takes to be analytical adequacy criteria for ethical theories: the principles of "deontic consequence" and "joint satisfiability". I am also going to discuss whether his formula-
tion of Collective Egoism yields the "right" prescriptions in its main test-case: Prisoners' Dilemma. As we shall see, it does not. Nor does it help to improve it. The improved version of Collective Egoism is able to deal with the two-person Prisoners' Dilemma, but it still misbehaves when we move to the many-persons cases. A certain type of "free rider"-problems proves to be especially troublesome.

In my presentation of Collective Egoism, I am going to make use of the paper by Howard Sobel (Sobel 1986), in which Österberg's position is both elucidated and criticized.¹

1. PRELIMINARIES

Before we continue any further, let us introduce the notion of a "situation" which plays an essential role in Österberg's arguments. A situation is defined by the (finite and non-empty) set $G$ of its agents, each of them having a non-empty but finite set of alternative actions at his disposal, the assignment of outcomes to every possible combination of individual actions, and the set of utility functions over outcomes, one function for each agent. So far, a situation is nothing but a "game", as games are understood in game-theory. However, if I understand Österberg correctly, the description of a situation contains more than that. For each subgroup $A$ of $G$, on each assumption about how $A$ might act, and for each subset $B$ of agents in $A$ and each possible combination of their actions, it is determined what the other agents in $G$ would have done, if this action-combination had been performed. As a consequence, it is determined what the outcome would have been.

The egoistic theories that we are here going to consider, are all "objective", in the sense that what is relevant to the normative status of the alternative actions are their actual rather than expected effects. What matters is the utility of the resulting outcome, not the expected utility. We do not need to consider the agents' expectations and probability assignments to other agents' behaviour.

While in this sense objective, the theories in question are "subjective" insofar as an agent's utility is to be understood as preference-satisfaction. His utility function over outcomes is supposed to reflect the degree to which the different outcomes satisfy his preferences² (even though the relevant preferences need not be the ones that the agent actually has but rather those that he would have if he were well-informed about what different outcomes involve and considered the matter carefully).