

A Hobbesian Derivation of the Principle of Universalization

Abstract. In this article, I derive a weak version of Kant's categorical imperative using an informal game-theoretic model. More specifically, I argue that Hobbesian agents would choose what I call the *weak principle of universalization*, if they had to decide on a rule of conflict resolution in an idealized but empirically defensible hypothetical decision situation. The discussion clarifies (i) the rationality requirements imposed on agents, (ii) the empirical conditions assumed to warrant the conclusion, and (iii) the political institutions that are necessary to implement the derived principle. The analysis demonstrates the moral significance of the weak principle of universalization and its epistemic advantage over the categorical imperative.