Abstract

We consider the one-shot Prisoner's Dilemma played by programs or machines, and show that the mutual cooperation is rather an ordinary event under the bounded rationality expressed by the computability. The *kin recognition player* (KRP, for short) is a program with the ability to recognize the opponent, and cooperate if and only if the opponent is *kin* to itself. We prove the existence of the KRP, and also of altruistic players which unilaterally self-sacrifice to the opponents that are kin to a reference KRP. It turns out that while any KRP is evolutionary stable, the self-sacrificing altruistic player is not.