

Towards a Formal Definition of Social Dilemma

What is a social dilemma in formal terms?

A non-cooperative normal form game is a social dilemma if there is a rationalizable pure strategy profile which is (weakly) Pareto inefficient.

Marcia L. Baptista¹

Pedro A. Santos^{2,3}

¹Instituto Superior Técnico,
Universidade Técnica de Lisboa

²CEAF at Instituto Superior Técnico,
Universidade Técnica de Lisboa

³GAIPS,
INESC-ID

[Problem]

Despite the importance of social dilemmas, a single **unified** and well-accepted **formal definition** of social dilemma has yet to be presented.

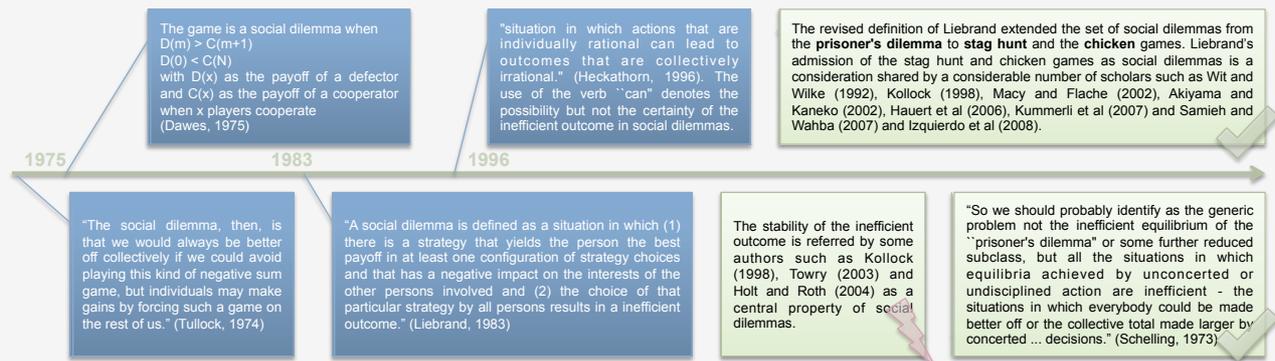
[Proposal]

We contend that a non-cooperative normal form game is a social dilemma if there is a rationalizable pure strategy profile which is (weakly) Pareto inefficient. By proposing this definition we are applying the **broadest** and most well-accepted **notion of collective rationality** and the **less stringent requirement** for the exercise of **individual rationality**.

[Implications]

With this definition we argue that social dilemmas should be perceived as **social settings prone to conflict** instead of settings in which conflict is certain. Furthermore, we hold that situations in which **long-term social inefficiency can arise** in the pursuit of the **short-term best interest of both the individual and society** should be considered social dilemmas.

The Definition of Social Dilemma Through Time



Difficulty in finding a formal definition for social dilemma stems from the fact that, even though it is commonly accepted that a social dilemma arises when there can be a **conflict between individual and collective interest**, authors seem to differ in the **formalization of the notions of individual and collective interest**.

Formalizing the concept of social dilemma

