

Learning to Resolve Social Dilemmas: A Survey (Abstract Reprint)

Shaheen Fatima¹, Nicholas R Jennings¹ and Michael Wooldridge²

¹ Loughborough University, UK

² Oxford University, UK

s.s.fatima@lboro.ac.uk, n.r.jennings@lboro.ac.uk, mjw@cs.ox.ac.uk

Abstract Reprint. This is an abstract reprint of a journal by [Fatima *et al.*, 2024].

Abstract

Social dilemmas are situations of inter-dependent decision making in which individual rationality can lead to outcomes with poor social qualities. The ubiquity of social dilemmas in social, biological, and computational systems has generated substantial research across these diverse disciplines into the study of mechanisms for avoiding deficient out-comes by promoting and maintaining mutual cooperation. Much of this research is focused on studying how individuals faced with a dilemma can learn to cooperate by adapting their behaviours according to their past experience. In particular, three types of learning approaches have been studied: *evolutionary game-theoretic learning*, *reinforcement learning*, and *best-response learning*. This article is a comprehensive integrated survey of these learning approaches in the context of dilemma games. We formally introduce dilemma games and their inherent challenges. We then outline the three learning approaches and, for each approach, provide a survey of the solutions proposed for dilemma resolution. Finally, we provide a comparative summary and discuss directions in which further research is needed

References

[Fatima *et al.*, 2024] Shaheen Fatima, Nicholas R. Jennings, and Michael J. Wooldridge. Learning to resolve social dilemmas: A survey. *J. Artif. Intell. Res.*, 79:895–969, 2024.