

# Transfer of learning about ToM opponents in zero sum games

Ismail Guennouni

December 9, 2019

## Introduction

We are interested in the way in which people build and use models of their opponent, when engaged in situations involving an interaction with strategic considerations. These situations arise frequently such as in negotiations, auctions, strategic planning and all other domains in which theory of mind abilities (Premack Woodruff, 1978) play a role in determining human behaviour. It is generally useful to reduce the complexity of real life phenomena we want to study using a simple model of reality that allows us to probe the interesting research questions in an experimental paradigm that reduces non relevant confounding effects and the complexity of real interactions. In the context of inter-dependent decision making, the study of simple games as a means of a model of more complex interactions played this essential role.

In order to explore learning in strategic interactions, it is essential to study repeated games in which players interact repeatedly and have the ability to learn about the opponent's strategies and preferences (Mertens, 1990). These consist of a repetition of a simple one-shot game such as the prisoner's dilemma (known as the base game) either with a finite horizon (known number of repetitions at inception) or 'infinitely' (meaning the number of repetitions is very high or unknown to the players). This framework introduces a time dimension that is essential to learning. It allows the study of whether and how a player takes into consideration, over time, the impact of its current and future actions on the future actions of the opponent and the future cumulative rewards.

The field of behavioural game theory was developed and contributed to a better understanding of the ways in which human players actually behave in experimental settings. A major contribution that preceded the development of behavioural game theory was the idea that humans possess a "bounded rationality". Simon (1972) explains that humans have limited cognitive capacities and as such cannot be expected to solve computationally intractable problems such as solving for Nash equilibria. Instead, they will try to 'satisfice' by choosing a strategy that is adequate in a simplified model of the environment, rather than an optimal one. This concept finds its natural application in 'level-k' theory, which posits that deviations from Nash equilibrium solutions are explained by the fact that humans have a heterogeneous degree of strategic sophistication. At the bottom of the ladder, level-0 players

are non-strategic and play either randomly or use a salient strategy in the game environment (Arad Rubinstein, 2012). Level-1 players are next up the ladder of strategic sophistication and will assume all their opponents belong to the level-0 category and as such will best respond to them given this assumption. Likewise, a level-2 player will choose actions that are the best response given the belief that all opponents are exactly one level below and so on

## **Methods**

## **Results**

## **Discussion**

## **Conclusion**