Adaptive Agents in Coalition Formation Games

Alex K. Chavez

University of Pennsylvania, Philadelphia, PA 19004, USA,
achavez@sas.upenn.edu

Abstract. Coalition formation games form an important subclass of mixed-motive strategic situations, in which players must negotiate competitively to secure contracts. This paper compares the performance of two learning mechanisms, reinforcement learning and counterfactual reasoning, for modeling play in such games. Previous work [CK04] found that while the former type of agent converged to theoretical solutions, they did so much more slowly than human subjects. The present work addresses this issue by allowing agents to update extensively based on counterfactual reasoning.

1 Introduction

Traditional solution concepts for n-person characteristic function games give stable sets of payoff assignments under certain rationality assumptions, but they have little to say about the process by which players arrive at these payoff sets. Furthermore, while it is possible that players realize these stable sets via deliberation, early experimental evidence [KR74] suggested that players engage in a learning process over repeated trials.

An explicit account of such a learning process can help to elucidate these problems. Modeling such a process, however, is complicated by the very large number of actions players must choose from. For example, players decide whether to accept coalition proposals, presumably based on the offer amount (drawn from an interval of the rational or real numbers), the offerer, and the other members of the coalition; they also decide whom to propose coalitions to, and how much to offer each member in the proposed coalition. To address these issues, Dworman, Kimbrough & Laing used genetic programming over the space of 1) the set of players, 2) lower and upper bounds of acceptance values, and 3) offer amounts.[1] They found that when mean payoffs of their system had settled, agents' payoffs approximated the game's quota values (a solution concept to be discussed).

While population models can be useful for exploring issues in agent-based systems and for representing evolutionary games, their interpretation for interactions on a smaller time scale (such as those of human players in repeated

[1] [DKL95a], [DKL95b], [DKL96], and [DKL96]

trials) is limited. Firstly, the timescales are different, and learning tends to be too slow. Secondly, agents in population models cannot be considered to be "minimally rational," as they are usually bare strategies (e.g., in replicator dynamics). Thus, it may be important to model human players as (boundedly) rational agents.

Borrowing from the approach of [ER98] in normal form games, [CK04] applied an individual learning model to five 3-person coalition games of [KR74]. They used a simple linear updating scheme (of reinforcement learning, see Equation 2) to update "aspiration levels" [MF02] of agents. These values, which remain within payoff bounds, represent the payoffs each agent believes it can get from other agents in a given coalition. Aspirations are mapped into behaviors by either a "greedy" rule, under which the agent maximizes over aspirations to the other agents, or a "matching" rule, under which the agent chooses probabilistically in ratios that match aspirations. [CK04] found that agents using the greedy rule converged to quota solutions in all five games, while matching agents never did. However, convergence took much longer for agents compared to human subjects in [KR74], who reached quotas in about four trials. This difference may have been due to calibration by subjects during practice rounds for which [KR74] did not report data, to deductive or counterfactual reasoning (reinforcement learning only uses information about received payoffs), or to experimental manipulations of a communication condition.

This paper compares the performance of a reinforcement learning model to that of one which incorporates information about payoffs *not* received, as well as payoffs received. It also examines the performance of agents utilizing a variety of offer behaviors. The remainder of the paper is organized as follows. Section 2 reviews coalition formation and presents the games studied in this paper. Section 3 discusses the learning model, updating equations, and offer behaviors. Section 4 compares the model's performance against theoretical solutions and again human data. Section 5 concludes.

2 Background and Description

2.1 Coalition Formation

Coalition formation describes a situation in which n players negotiate competitively to secure advantageous contracts. Each player can act alone to attain some fixed payoff. However, by pooling resources with one or more other players and forming a *coalition*, a player can exploit its comparative advantages and thereby negotiate for payoffs that are larger than the sum of the payoffs the individuals in the coalition can get by acting alone.[2] The pay-

[2] This property is known as superadditivity, and holds for all games studied here. The payoff for acting alone is normalized to zero, and the assumption is made that all coalitions have payoffs greater than or equal to zero. Payoffs are assumed to be on the same scale, or in units of "transferable utility" [LR57].