A Note on Working Memory in Agent Learning

Fang Zhong

Georgia Institute of Technology, Atlanta, GA, USA,
fang.zhong@mgt.gatech.edu

Abstract. An important dimension of system and mechanism design, working memory, has been paid insufficient attention by scholars. Existing literature reports mixed findings on the effects of the amount of working memory on system efficiency. In this note, we investigate this relationship with a computational approach. We design an intelligent agent system in which three agents, one buyer and two bidders, play an Exchange Game repeatedly. The buyer agent decides whether to list a request for proposal, while the bidders bid for it independently. Only one bidder can win on a given round of play. Once the winning bidder is chosen by the buyer, a transaction takes place. The two parties of the trade can either cooperate or defect at this point. The decisions are made simultaneously and the payoffs essentially follow the Prisoner's Dilemma game. We find that the relationship between working memory and the efficiency of the system has an inverted U-shape, i.e., there seems to be an optimal memory size. When we mixed agents with different memory sizes together, agents with the same amount of working memory generate the most efficient outcome in terms of total payoffs.

1 Introduction

One of the enduring challenges in electronic commerce is modeling and implementing formal conversations among agents [KL97]. In facing this challenge, researchers need to deal with the design of intelligent agents who can not only communicate with each other effectively, but also respond promptly, efficiently, and intelligently in dynamic environments. Previous research has shown the potential of intelligent agents with learning abilities for achieving cooperation or a high level of trust in trading environments without legal enforcement.[1] The focus of these studies is on applying Q-learning (a form of reinforcement learning[2]) to the design of agents. Learning, however, is only one aspect of intelligence. Another important dimension is the working memory of an agent, i.e., how much historical information is incorporated by an agent in its decision-making processes.

This design question is not trivial. Sandholm and Crites[3] conducted a study in which two artificial agents play the Iterated Prisoner's Dilemma

[1] [KWZ02,SC95,ZKW02]
[2] [KLM96,SB98,WD92]
[3] [SC95]

(IPD) game. The two agents could be either Q-learners or TIT-FOR-TAT strategy players. They showed that agents with larger memories seemed to be able to gain larger total payoffs when playing against agents with smaller memories. However, the focus of their paper is not on the influence of working memory, but on applying Q-learning in multi-agent system design. One major difficulty of expanding memory size in the context of reinforcement learning is the resulting explosive state space [KLK04]. Due to the inherent computational complexity of large-scale investigation in this topic, we first need to find evidence in and results for a relatively small system. That is the purpose of this note.

In our computational approach, learning agents with different memory sizes play a multi-stage trading game repeatedly. The last stage of the game is essentially a simultaneous Prisoner's Dilemma game. Our results show that the impact of memory size on agent performance has an inverted U-shape. There seems to be an optimal memory size. Specifically, it appears that for a two-by-two game, conditioning the bidder's decision on information from the last two periods is optimal compared to memory sizes of one or three periods. Moreover, when agents with different memory sizes trade with each other, having more memory does not pay off. These findings contravene the results reported in Sandholm and Crites's paper [SC95].

Another contribution of this note is the provision of a model in which agents have the option of choosing their trading partners. In social exchanges involving trust, participants not only decide how to bargain with one another; they also choose with whom to bargain. Most studies focus on the bargaining or exchange process itself, with participants either fixed or matched randomly against one another.[4] However, the situation in which a potential participant in an exchange can choose a partner, or even whether to enter the exchange at all, has rarely been investigated. It is important to understand, in this more realistic context, whether intelligent agents can make the right choices when facing different potential partners.

The rest of this note is organized as follows. §2 provides motivation for and description of the Exchange Game we study throughout the note. §3 describes briefly the design of the agents. The experimental design and results are discussed in §4, after which we conclude the note.

2 The Exchange Game

The model proposed here is motivated by the type of exchange existing in online markets, in which a buyer initiates an auction by posting a Request for Proposal (RFP). Listing an RFP is optional. The buyer incurs this cost only if it decides to solicit proposals. This cost comes from preparing the RFP and posting it in the right market. Bidders subsequently submit bids. To consider the endogenous entry decision of bidders, we include bidding cost

[4] E.g., [CB98], [HW98], [SC95], and [ZKW02].