Learning and Tacit Collusion by Artificial Agents in Cournot Duopoly Games

Steven O. Kimbrough[1], Ming Lu[2], and Frederic Murphy[3]

[1] University of Pennsylvania, Philadelphia, PA, USA,
kimbrough@wharton.upenn.edu
[2] University of Pennsylvania, Philadelphia, PA, USA,
milu@wharton.upenn.edu
[3] Temple University, Philadelphia, PA, USA,
fmurphy@temple.edu

Abstract. We examine learning by artificial agents in repeated play of Cournot duopoly games. Our learning model is simple and cognitively realistic. The model departs from standard reinforcement learning models, as applied to agents in games, in that it credits the agent with a form of conceptual ascent, whereby the agent is able to learn from a consideration set of strategies spanning more than one period of play. The resulting behavior is markedly different from behavior predicted by classical economics for the single-shot (unrepeated) Cournot duopoly game. In repeated play under our learning regime, agents are able to arrive at a tacit form of collusion and set production levels near to those for a monopolist. We note that Cournot duopoly games are reasonable approximations for many real-world arrangements, including hourly spot markets for electricity.

1 Introduction

Strategic behavior by rationally limited agents is interesting for a number of reasons. We draw the reader's attention to two in particular. First, *real* agents are rationally limited. The rationally ideal agents of economics and classical game theory are abstractions, idealizations created for the legitimate purpose of tractable modeling. Without questioning the legitimacy of this, there remains the question of how systems of non-ideal agents will behave. Even if we think that humans in standard market conditions do indeed approximate the rational ideal, no one maintains that birds, bees, monkeys up in trees, and artificial agents do so. And they are interesting, too. Second, in many contexts formal game theory makes predictions that are in various ways unsatisfactory, or unsatisfying, even for rationally ideal agents. ([Col95] and [Kre90] are excellent and accessible reviews of this broadly-accepted assertion.) A case in point occurs when, as is generally the case in repeated games, the number of equilibria in the super-game (the game consisting of repetitions of a sub-game) is large or even infinite. Predicting that the outcome of the super-game will occur at some equilibrium is unsatisfying because it is so

unspecific. Further, classical game theory has little to say about the process of playing and finding good strategies, or how to implement effective agents in strategic contexts, a point developed in [DKL96].

The study of strategic behavior by artificial agents offers a means of addressing many of these and other issues. Methodologically, we may call this a form of *algorithmic game theory,* complementing classical, or *a priori*, game theory, and behavioral game theory (aka: experimental economics) [KL04]. In what follows we present results from this perspective. We examine play and learning by artificial agents in the context of a *repeated* Cournot duopoly game [Cou97]. In §2 we describe the particular version of the model used in our study. §3 discusses briefly the use and appropriateness of the Cournot model in modern electricity markets, in which repetition of play occurs on an hourly basis, which means 8760 plays per year. §§4–5 discuss our learning model for agents in repeated Cournot games. The model is simple and cognitively realistic, yet it departs in an important way from the literature on learning by artificial agents in games. We present our results in §6 and conclude with a discussion in §7.

2 The Duopoly Game: Holt's Cournot Model

The Cournot duopoly model is a staple of classical economics and its textbooks (see, e.g., [Var03], but any standard microeconomics text will do). Because it has been studied both in the laboratory with human subjects (by Holt [Hol85]), and with simple reinforcement learning agents (by Kimbrough and Lu [KL03]), and because it is unproblematically representative of Cournot duopoly models generally, we report on agent learning behavior in the context of a very simple Cournot model, as follows.

There is a duopoly in which the two competing firms are the players and produce a homogeneous product. They individually decide only their individual levels of production, x_1 and x_2. Variable costs are 0 and demand is linear. The total price paid for the joint production, P, is a linear function of the total output:

$$P = A - B(x_1 + x_2) \tag{1}$$

(It is assumed that all variables are greater than 0.) The profit for firm i, $\pi(x_i, x_{-i})$, is a function of both its production, x_i, and the production of the other firm, x_{-i}. (We use conventional notation: $-i$ denotes the player or players in the game *other than i*. This amounts to

$$\pi(x_1, x_2) = x_1[A - B(x_1 + x_2)] \tag{2}$$

for firm 1 in the case, as we have here, of two players. Similarly, the profit for firm 2 is

$$\pi(x_2, x_1) = x_2[A - B(x_1 + x_2)] \tag{3}$$