

Conducting Experiments with the Axelrod Tournament Demo

References for Team Exercise 5:

- [1] ** Chris Cook, Axelrod Tournament Demo. Home Page (including a link to automatic installation software) is available ON-LINE at

<http://www.econ.iastate.edu/tesfatsi/demos/axelrod/axelrodt.htm>

- [2] ** Leigh Tesfatsion, “Notes on Axelrod’s Iterated Prisoner’s Dilemma (IPD) Tournaments,” ON-LINE (Syllabus Section IV.A)

- [3] ** Leigh Tesfatsion, “Experimental Design: Basic Concepts and Terminology,” ON-LINE (Syllabus Section I.A)

In 1979, Robert Axelrod (University of Michigan) posed an intriguing question about the Iterated Prisoner’s Dilemma (IPD) game: What type of strategy (if any) *ensures* good individual performance over the long haul when one is engaging in IPD game play in round-robin fashion with multiple “strangers” whose IPD strategies are not known in advance? As discussed in class, Axelrod decided to explore this question by running a computer tournament. Various professional game theorists were solicited for IPD strategies in the form of computer algorithms, and the submitted strategies were then pitted against each other in round-robin IPD game play. The winner of this tournament was the simple IPD strategy *Tit-for-Tat (TFT)* defined as follows: *Start by cooperating, then do whatever your rival did in the previous iteration.* Axelrod later conducted a second larger tournament, together with derived “ecological” tournaments, all of which Tit-for-Tat again won.

The Axelrod Tournament Demo developed by Chris Cook is available at

<http://www.econ.iastate.edu/tesfatsi/demos/axelrod/axelrodt.htm>

This demo captures salient aspects of Axelrod’s 1979 tournament. The user can specify as treatment factors the types of agents (strategies) to be entered into the tournament along with the number of agents of each type. The user can also specify the number of iterations for each tournament. In addition, however, the demo extends the original Axelrod Tournament by permitting the user to specify the type of 2×2 game the agents engage in. Specifically, the user can specify customized 2×2 payoff matrices for the four possible move combinations (CC,CD,DC,DD) available to the players, and the user can freely associate these customized payoff matrices with different agent types. The user can then observe how well each type of agent fares over time as the agents participate in iterated round-robin game play.

This team exercise asks each team to use the Axelrod Tournament Demo to develop and implement an experimental design for exploring the extent to which TFT is successful (or not) in game contexts differing from the standard IPD game. **Three restrictions are imposed on your experimental design as follows – please make absolutely sure you satisfy all three restrictions so different team findings can be compared in an interesting way:**

Restriction 1: At the end of this exercise are two tables containing two payoff matrices with specific numerical payoffs, the first for a specific CHICKEN GAME and the second for a specific STAG HUNT GAME. *Even-numbered* teams are asked to assign the specific numerical Chicken Game payoffs in Table 1 to each agent they enter into their tournament. *Odd-numbered* teams are asked to assign the specific numerical Stag Hunt Game payoffs in Table 2 to each agent they enter into their tournament.

Restriction 2: Be sure to include TFT agents within your tournament agent populations, so that their relative performance can be evaluated in this non-IPD tournament context.

Restriction 3: Be sure to include at least one RANDOM agent within your tournament agent populations, as did Axelrod in his original Tournament, as a benchmark case of a non-reactive agent without a bias towards being nice or nasty.

Part A: Formulate an Interesting Hypothesis

Let the performance of each agent type in each tournament run be evaluated in accordance with its average payoff per game played, as measured by the “Average Payoff” column reported in the screen accessed through the “Agent Stats” tab on the main Demo screen.

Referring to reading [3] above, carefully formulate an hypothesis (conjecture) that your team judges to be interesting and substantial regarding how a **systematic change** in some chosen treatment factor OTHER THAN PAYOFFS might affect the performance of the agent types in your tournament (including, in particular, the performance of TFT agents).

IMPORTANT QUALIFICATIONS: Recall that the seed value for the pseudo-random number generator CANNOT be considered a “treatment factor.” As clarified below, these seed values are to be used to generate sample runs for each particular treatment specification you wish to test.

Part B: Explore Your Hypothesis within an Experimental Design

Use the Demo to experimentally explore the validity of the hypothesis you proposed in Part A. Specifically:

1. Choose a range of values (at least three) to be tested for your chosen treatment factor in Part A.
2. Set fixed values for all OTHER structural features of the Demo, to be retained throughout all experimental runs, and report these fixed values along with the type of game (Chicken or Stag Hunt) that you have been assigned.
3. For each value of your treatment factor to be tested, use the Demo to conduct N runs ($N \geq 10$) using N distinct seed values for the pseudo-random number generator.
4. For each value of your treatment factor to be tested, and for each run $n = 1, \dots, N$ conducted for this treatment factor value, report:
 - (a) the value of the treatment factor that is being tested;
 - (b) the seed value (the identifier for the run);
 - (c) the performance (average payoff per game played) attained by each agent type participating in this run.
5. For each tested treatment factor value, report the mean and standard deviation for the performance of each agent type across the N runs you conducted in Step 4 using the performance findings you reported in Step 4(c) for individual runs.
6. For each tested treatment factor value, examine the variation across the N runs in the performance of each agent type by clicking the “expand” feature on the “Graph” menu screen for each agent type. Report on any interesting observations here as well.

Part C (3 Points): Analysis of Findings

As best you can, provide an explanation and interpretation for the experimental performance findings you reported in Part B. Do these findings provide any support for the hypothesis you proposed in Part A? Or does your hypothesis in Part A appear to be inconsistent with these findings? In particular, what can you conclude about the relative performance of TFT agents in your assigned game? Justify your assertions carefully.

		PLAYER 2	
		C	D
PLAYER 1	C	(R=2,R=2)	(S=1,T=3)
	D	(T=3,S=1)	(P=0,P=0)

Table 1: A Chicken Game ($T > R > S > P$)

		PLAYER 2	
		C	D
PLAYER 1	C	(R=3,R=3)	(S=0,T=2)
	D	(T=2,S=0)	(P=1,P=1)

Table 2: A Stag Hunt Game ($R > T > P > S$)