Axelrod's Round-Robin Contest in a Classroom Setting

YAMAMOTO, Kazuya

Abstract

This study replicates Robert Axelrod's first round-robin tournament of the prisoner's dilemma, using strategies developed by members of the author's laboratory. The qualities of strategies between specialists (entrants in Axelrod's tournament) and beginners (participants in the contest reported in this paper) were not too different from one another. Unlike the results of Axelrod's tournament, TIT FOR TAT did not win our contest. However, the so-called *nice* strategies submitted in our contest remained as strong as those in Axelrod's contest. Overall, our results reveal that Axelrod's experiments produced robust results.

INTRODUCTION

The prisoner's dilemma has been a prominent focus of game theory. The equilibrium analysis of the game has found rational consequences that may be either optimistic or pessimistic, depending on the conditions imposed on each model. Another approach to the prisoner's dilemma is to reveal the outcome of the interactions between agents by relaxing rationality. Psychological experiments have placed people in a prisoner's dilemma so as to observe their reactions, many of which were conducted during the 1950s and 70s. For example, *the Journal of Conflict Resolution* published articles in the early period reporting experimental results regarding prisoner's dilemmas and other games (e.g., Radlow, 1965; Conrath, 1970). Robert Axelrod proposed another method in the late 1970s — his idea was to allow agents to play the prisoner's dilemma on a computer instead of playing it in a laboratory. His round-robin computer tournaments of the prisoner's dilemma impacted all of the social sciences. The two reports that discussed the results of his contests are among the most read articles in the social sciences published in the 1980s (Axelrod, 1980a, 1980b, 1984).

This study replicates the first contest of Axelrod's two tournaments. The experiment was conducted in the author's class in January of 2023.(1) Students wrote about their original strategies and submitted their programs to the tournament. Many students were beginners in computer programming languages and were studying Axelrod's research for the first time.

This study has two aims. First, it compares the strategies submitted to Axelrod's tournament with those used in our contest. The experiment in this study reveals how beginners in both topics (i.e., programming and the prisoner's dilemma) tend to develop strategies to win a contest of the iterated prisoner's dilemma. Meanwhile, the entrants in Axelrod's first tournament were specialists in such experiments, as exemplified by Anatol Rapoport, a psychologist and winner of the contest. Comparing the differences between the two sets of strategies will provide knowledge on how ordinary people are likely to develop strategies if they are placed in situations represented by these contests.

The second purpose is to examine whether the results of Axelrod's two experiments are robust. The implications of his experiments are prevalent among both scholars and practitioners. However, these implications are highly dependent on the specific strategies used in competitions. If a different set of strategies developed by non-specialists yields the same results, the implications of Axelrod's tournaments would be

(1) Kanta Kishida, Shota Maeno, Kazuma Nagai, Satoru Nakano, Ryo Noguchi. Namiho Okuno, Yuhi Sonehara, Fuwa Tomita, Tasuku Yakura, and Hirai Yokota are the entrants for the contest reported in this paper.

relevant to a broader swathe of social situations.

DESIGN OF THE CONTESTS

In both Axelrod and our own tournaments, the entrants submitted the strategy that they considered to be the best performing.(2) The tournament followed a round-robin format, with each pair of strategies competing in a game in which the prisoner's dilemma was repeated 200 times. Individual strategies also faced itself in the round-robin. Five round-robins were executed in Axelrod's contest, whereas we ran the tournament thirty times in our experiment. Multiple tournament runs allow for variances caused by strategies using probability calculations. The tournaments used the payoff matrix shown in Fig. 1.

		Player B	
		Cooperation	Defection
Player A	Cooperation	(3, 3)	(0, 5)
	Defection	(5, 0)	(1, 1)

Fig.1　The Prisoner's Dilemma

Thirteen strategies were submitted to our contest. Additionally, the RANDOM and TIT FOR TAT (TFT) strategies were included in advance. Each submitted strategy is as follows:

CTRG: This strategy starts with the choice of cooperation, but the number of consecutive defections increases proportionally with the number of defections taken by the other player. For example, for the other player's first defection, CTRG will defect in the next move. However, if CTRG finds that the other player has returned to cooperative behavior in that move, CTRG will also return to a cooperative choice for the following move. If the other player defects at some later move again, CTRG defects twice during the following moves despite the other player's return to cooperation immediately after its second defection. Thus, CTRG's memory is cumulative. CTRG never returns to cooperation if the number of the other player's defections amounts to ten.

TTRG: This strategy judges the other player by referring to the other player's behavior during the first 20 moves. TTRG begins with a cooperative choice. Once the other player defects at any move before the twentieth move, TTRG continues to defect thereafter. However, if the other player does not defect in the first 20 moves, TTRG continues to cooperate until the end of the game despite the other player's defection in the following moves.

HANTEI: This strategy is a variant of the TFT strategy, but incorporates a mechanism that identifies two specific strategies. One such strategy is RANDOM. If HANTEI judges the other player as following the RANDOM strategy, it continues to defect thereafter. Another is the one that adopts TFT but sometimes defects in order to exploit the other player. If HANTEI identifies the other player as this type, it chooses not to defect but to cooperate against the other player's defection and induces the other player to return to cooperation. Because the main rules of both strategies are based on TFT, HANTEI's cooperative choice against the other player's defection allows them to return to reciprocal cooperation, although the other player might defect again during later moves.

PAST: This strategy values the gains obtained from recent moves. Starting with defection for the first three moves, it decides on the following choices by referring to the payoffs gained in the latest three moves. If the sum of the gains in the preceding three moves is either zero or 15, it chooses to defect. If the sum is 2, …, 5, 7, or 9, it cooperates with a probability of 0.7 and defect with a probability of 0.3. If the sum is 1, 6, 8, 10, 11, or 13, the probability of cooperation is 0.3 while that of defection is 0.7.

RND135: This strategy involves a mixture of RANDOM and TFT. It consists of six moves. It chooses cooperation with a probability of 0.1 and defection with a probability of 0.9 at the first move. It employs TFT during the second move. The third move uses a random strategy again, but cooperation is selected with a probability of 0.3 and defection is chosen with a probability of 0.7. TFT is used again during the fourth move. One of

(2) In our contest, participants were allowed to submit two or more strategies.

the two choices is equally probable for the fifth move. TFT is used for the sixth move, before returning to the beginning of the cycle.

SC: This is a variant of TFT following a cycle of five moves. It cooperates until the fifth move, then counts the number of cooperations performed by the other player during this period. If the number is at least three, SC continues to cooperate during the next five moves. Otherwise, SC defects during the next five moves. The same cycle of evaluation continues until the end of the game.

KNG and SPY: These two strategies collude with one another.(3) Starting with mutual cooperation, the two strategies share information regarding the sequence of their choices until the tenth move in advance. Using this information, they judge whether the present competitor is their ally after the tenth move ends. If KNG identifies another player as SPY, it chooses to defect thereafter. In all other situations, TFT was used. Meanwhile, if SPY judges the other player as following the KNG strategy, it chooses to cooperate until the end of the game. Otherwise, SPY continues to defect until the end of the game. Thus, SPY's dedication to KNG allows the latter to maximize its payoffs in the match between KNG and SPY.

TCTD: This strategy considers a three-move cycle. It randomly determines choices during the first three moves. If it finds that the other payer has cooperated two or more times during that cycle, it defects twice in the next cycle. If the other player defects two or more times, TCTD reacts randomly in the next cycle. These cycles continue until the end of the game.

TFTD: This strategy begins with a cooperative choice. It then uses RANDOM with a probability of 0.1 while employing TFT with a probability of 0.9.

DWC: This strategy adopts an approach similar to DOWNING, which is a strategy used in the original tournament. It calculates two probabilities: the probability of the other's cooperation after the DWC's cooperation, and the probability of the other's cooperation after the DWC's defection. It then calculates the expected values of DWC's cooperative and defect choices using the two probabilities. The choice that produces a greater value is selected as DWC's next move. The calculations of the expected values are mainly based on a specific number of latest moves; however, DWC also refers to older records by discounting them. Unlike DOWNING, DWC begins with a cooperative choice.

TYPES OF STRATEGIES

Table 1 categorizes the entries in our tournament into different types of strategies and indicates whether they are *nice*. A *nice* strategy never defects before the other player does so first; this property is a key idea in Axelrod's interpretation of his tournaments. Because TFT is the most investigated strategy in studies of the repeated prisoner's dilemma game and Axelrod's tournaments, many entries in our contest also incorporate this rule into their programs. HANTEI and SC are straightforward expansions of TFT. KNG is also a variant of TFT. There are entries that combine TFT with other well-known strategies. CTRG is a variant of the trigger strategy; however, it incorporates TFT's reciprocity into its algorithm. RND135 and TFTD combine RANDOM with TFT. The remaining entries did not use this reciprocal principle. TTRG simply modifies the trigger strategy. PAST only considers past

Table 1 Types of Strategies

	Type	Nice
CTRG	Trigger/TFT	Yes
TTRG	Trigger	Yes
HANTEI	TFT	Yes
PAST	Other	No
RND135	RND/TFT	No
SC	TFT	Yes
KNG	TFT	No
SPY	All-C(D)	No
TCTD	Other	No
TFTD	RND/TFT	No
DWC	Other	Yes

NOTE: DWC might be either *nice* or not *nice* depending on the setting of parameter values. The entrant submitted a program that lets DWC act nicely. TFT and RND are not included in the table.

(3) As multiple entries were not prohibited in our contest, a participant submitted two strategies that form an alliance. A similar strategy, which is called master and slave, was developed for the 20th anniversary competition held in 2004 (Osawa and Imai, 2007).

Table 2　The Result of Our Contest

	Other Players													
Player	DWC	HANTEI	SC	TTRG	TFT	CTRG	KNG	RND135	TFTD	SPY	RND	TCTD	PAST	*Average Score*
DWC	600	600	600	600	600	600	556	385	538	201	571	558	585	538
HANTEI	600	600	600	600	600	600	589	510	396	205	550	544	536	533
SC	600	600	600	600	600	600	593	329	508	212	446	446	436	505
TTRG	600	600	600	600	600	600	221	321	262	221	594	602	587	492
TFT	600	600	600	600	600	600	498	530	239	216	448	419	436	491
CTRG	600	600	600	600	600	600	232	346	241	216	582	583	576	490
KNG	546	604	593	201	503	232	212	531	243	968	448	434	434	457
RND135	673	522	235	174	530	219	524	352	299	185	486	435	493	394
TFTD	635	440	603	273	244	244	247	313	222	216	467	432	461	369
SPY	276	250	227	201	221	221	23	338	218	590	547	573	575	327
RND	176	216	455	108	450	139	439	378	419	210	450	375	433	326
TCTD	179	208	373	105	423	130	407	370	394	146	486	414	494	317
PAST	111	216	395	106	439	130	444	375	409	125	458	372	443	309

NOTE: The values are rounded down to the nearest unit. The line between CTRG and KNG indicates that all *nice* strategies rank higher than all strategies that are not *nice*.

gains in deciding the next choice. Instead of choosing cooperation, TCTD responds to the other player's cooperation through defection. SPY follows either an all-cooperation or all-defection strategy after the tenth move. The calculations employed by DWC do not ensure reciprocal responses to the other player.

Axelrod's (1980a) appendix explains the rules of the 15 strategies used in his first tournament. The participants in our contest read Axelrod (1984), but were unlikely to have read Axelrod (1980a). Therefore, they did not know the details of the programs submitted to Axelrod's tournament, except for a brief explanation of several rules described in Axelrod (1984).

Overall, the rules implemented by our entrants were simpler than those implemented in Axelrod tournament. Some entrants in Axelrod's tournament submitted extremely complicated programs — such as STEIN and RAPOPORT, which used chi-squared tests — while no such rule was submitted to our contest. Nevertheless, many entrants in either Axelrod or our tournaments present similar ideas. Two strategies in Axelrod's contest — TIDEMAN and CHIERUZZI, and SHUBIK — and one strategy in our contest, CTRG, respond similarly to the other player by increasing in their frequency of punishment for the other player's cumulative defections. NYDEGGER in Axelrod's contest and PAST in our contest are similar in that both strategies calculate scores using the latest three outcomes of moves and determine the next choice. Both Axelrod's FRIEDMAN and our TTRG adopt a trigger strategy, although TTRG's rule is more complicated than that of FRIEDMAN.

Thus, the strategies formed by beginners are not very different from those developed by specialists. This observation indicates that the situation formed by Axelrod's first-round tournament is not that of a peculiar environment comprising *odd* strategies provided by specialists, and that his competition represents a common situation.

THE RESULT OF THE CONTEST AND DISCUSSION

Table 2 presents the tournament results.[4] TFT ranked fifth. Axelrod held two tournament sessions and TFT won twice. Our results indicate that the dominance of TFT, as presented by Axelrod's experiments, is not universal but depends on situations formed by specific entries in the contests. Second, the strength of *nice* strategies observed in Axelrod's tournaments was also confirmed by our contest. We have six nice strategies,

(4) Our contest was arranged in a classroom with the purpose of education as well as research. The author reviewed the entered programs submitted by the students and corrected mistakes. Even so, there might remain bugs in programs. Thus, implications derived from this report are tentatively presented, although we are fairly confident in the correctness of our execution.

including TFT. As Table 2 indicates, these strategies dominated the other strategies. Since these six strategies are nice, they never defect against each other and thereby gain 600 payoffs in each pair of nice strategies.

We should note that two of the six strategies, DWC and TTRG, are nice but are not variants of TFT. The view that TFT is the best way to win negotiations or conflicts has prevailed since Axelrod's research was published. U.S. sanctions against other countries have sometimes been justified by resorting to this view. However, our results indicate that following a TFT strategy — or its variants (i.e., having the property of reciprocity) — is not sufficient for winning a tournament. Instead, being nice may have been sufficient. In fact, DOWNING begins with defection and ranked tenth among the 15 programs in Axelrod's first tournament. By contrast, DWC in our contest — a variant of DOWNING — won the contest by starting with cooperation.

Third, attempts to exploit other players were unsuccessful. TCTD is a strategy that implements such a rule. TCTD defects when the other player cooperates, thereby attempting to gain a greater payoff. However, it was ranked twelfth. Fourth, collusion was ineffective for winning the tournament because this action was only effective between the colluding players; KNG gained greatly from SPY, but it could not perform well against other strategies, probably due to the part of its program that aims to identify the other player. Finally, strategies that focused on their own payoffs or disregarded the responses of other players lost the tournament. PAST and strategies with random behavior are included in these strategies in our contest.

CONCLUDING REMARK

The results of Axelrod's contests have been interpreted broadly, and his experiments have been criticized — both reasonably and unreasonably. Nevertheless, his studies have been influential in both practice and research. Overall, the results of our experiment are congruent with those of Axelrod's tournaments. The robustness of the original experiments is confirmed in this report. This quality has allowed the research to be significant for 40 years, despite repeated criticisms.

REFERENCES

Axelrod, Robert. (1980a). "Effective Choice in the Prisoner's Dilemma," *Journal of Conflict Resolution*, 24 (1): 3–25.

Axelrod, Robert. (1980b). "More Effective Choice in the Prisoner's Dilemma," *Journal of Conflict Resolution*, 24 (3): 379–403.

Axelrod, Robert. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Conrath, David W. (1970). "Experience as a Factor in Experimental Gaming Behavior," *Journal of Conflict Resolution*, 14 (2): 195–202.

Osawa, Hirotaka and Imai Michita. (2007). "Ningendoshi niyoru Kurikaeshigata Shujin no Jirenma Gemu no Jisshi to Kekka," *Gemu Puroguramingu Wakushoppu 2007 Ronbunshu*, 12: 188–194.

Radlow, Robert. (1965). "An Experimental Study of 'Cooperation' in the Prisoner's Dilemma Game," *Journal of Conflict Resolution*, 9 (2): 221–227.