

NOTES

Chapter 1. *The Problem of Cooperation*

1. For useful illustrations of these applications to international politics, see the following sources: the security dilemma (Jervis 1978), arms competition and disarmament (Rapoport 1960), alliance competition (Snyder 1971), tariff negotiations (Evans 1971), taxation of multinational firms (Laver 1977), and communal conflict in Cyprus (Lumsden 1973).

2. The Prisoner's Dilemma game was invented in about 1950 by Merrill Flood and Melvin Dresher, and formalized by A. W. Tucker shortly thereafter.

3. The situations that involve more than pairwise interaction can be modeled with the more complex n -person Prisoner's Dilemma (Olson 1965; G. Hardin 1968; Schelling 1973; Dawes 1980; R. Hardin 1982). The principal application is to the provision of collective goods. It is possible that the results from pairwise interactions will help suggest how to undertake a deeper analysis of the n -person case as well, but that must wait. For a parallel treatment of the two-person and n -person cases, see Taylor (1976, pp. 29-62).

4. The value received from always defecting when the other is playing TIT FOR TAT is:

$$\begin{aligned} V(\text{ALL D}|\text{TFT}) &= T + wP + w^2P + w^3P \dots \\ &= T + wP(1 + w + w^2 \dots) \\ &= T + wP/(1-w). \end{aligned}$$

5. If the other player is using a strategy of permanent retaliation, you are better off always cooperating than ever defecting when $R/(1-w) > T + wP/(1-w)$ or $w > (T-R)/(T-P)$.

6. This means that the utilities need only be measured as an interval scale. Using an interval scale means that the representation of the payoffs may be altered with any positive linear transformation and still be the same, just as temperature is equivalent whether measured in Fahrenheit or Centigrade.

7. For the implications of not assuming deliberate choice in an evolutionary model of economic change, see Nelson and Winter (1982).

Notes to pages 30-56

Chapter 2. *The Success of TIT FOR TAT*

1. The second round of the tournament was described in the text.

2. This is a broader definition of forgiveness than Chammah (1965, pp. 72-73), which is the move after receiving the sucker's payoff, S.

3. In the five games between them, the average score was 230 for TAT and 230 for JOSS.

4. In the environment of the 15 rules of the tournament, TIT FOR TWO TATS averages 542 points. This compares to 504 points. TIT FOR TWO TATS averages 504 points. TIT FOR TWO TATS averages 504 points. LOOK AHEAD averages 520 points.

5. This probability of ending the game at the expected median length of a game would be 2.5. In a sample of 1000 games, the length of a game would be 2.5 on average. In a sample of 1000 games, the length of a game would be 2.5 on average. In a sample of 1000 games, the length of a game would be 2.5 on average.

6. This reproduction process creates a simulation in which the average score achieved by each rule with each of the rules, where the weights are the numbers of the other rules in the initial generation.

7. This simulation of future rounds of the tournament is a weighted average of the scores of a given rule in the current generation. The weights are the numbers of the other rules in the current generation. The numbers of a given rule in the next generation are proportional to the product of its numbers in the current generation. This procedure assumes a Markov process. It is the only instance in this book where a cardinal, rather than merely interval, interpretation is used.

Chapter 3. *The Chronology of the Game*

1. Those familiar with the concepts of game theory will recognize the definition of a collectively stable strategy as a strategy that is not invaded by any other strategy. My definitions of invasion and collective stability are based on Maynard Smith's (1974) definitions of invasion and collective stability. His definition of invasion allows a newcomer to do as well as a native meeting a native, providing the newcomer does better than a newcomer meeting a native. I have provided new definitions to simplify the proofs and to clarify the effect of a single mutant and the effect of a single rule which is evolutionarily stable is also collectively stable. The definitions which will never be the first to defect, the definitions in the text remain true if "evolutionary stability" is replaced by "collective stability" with the exception of the characterization of the character B, where the characterization is necessary but