THE ROYAL
SOCIETY

The efficiency of adapting aspiration levels

Martin Posch[1], **Alexander Pichler**[2] **and Karl Sigmund**[2,3]

[1]*Universität Wien, Institut für Medizinische Statistik, Schwarzspanierstrasse 17, 1090 Wien, Austria*
[2]*Universität Wien, Institut für Mathematik, Strudlhofgasse 4, 1090 Wien, Austria*
[3]*International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria*

Win–stay, lose–shift strategies in repeated games are based on an aspiration level. A move is repeated if and only if the outcome, in the previous round, was satisficing in the sense that the pay-off was at least as high as the aspiration level. We investigate the conditions under which adaptive mechanisms acting on the aspiration level (selection, for instance, or learning) can lead to an efficient outcome; in other words, when can satisficing become optimizing? Analytical results for 2×2 games are presented. They suggest that in a large variety of social interactions, self-centred rules (based uniquely on one's own pay-off) cannot suffice.

Keywords: games; satisficing; learning rules; natural selection

1. INTRODUCTION

In a game theory without rationality (see Rapoport 1984), players are not assumed to be able to fully understand the situation in which they are engaged. Their moves are based on knee-jerk rules rather than on strategic analysis. Possibly the simplest of such rules is the win–stay, lose–shift principle, which consists of repeating an action if it proved successful, and in switching to another action if not. Suppose that we were playing a machine with two levers, one resulting in a positive, the other in a negative outcome. The win–stay, lose–shift principle would result in our repeating the action with the positive outcome; if we erroneously tried the wrong action, we would switch back, in the next round, to the right action. Many experiments have shown that such a behaviour, or some approximation of it, is widespread among human and animal actors. Interestingly, this crudest form of a learning rule works even in situations involving several agents, as in the so-called minimal social situation (Colman 1995).

The win–stay, lose–shift principle was originally formulated by Thorndike:

> 'Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction are more firmly connected with the situation; those which are accompanied or closely followed by discomfort have their connection with the situation weakened.' (Thorndike 1911, p. 244)

The wide range of validity of this principle was soon recognized (see for example, Hoppe 1931; Rescorla & Wagner 1972). In the hands of H. Simon, satisfaction-seeking behaviour became a leading contender for explaining social and economic decision making (see Simon 1955, 1957, 1962; Winter 1971; Radner 1975). A considerable amount of empirical evidence suggests that the behaviour of individuals and firms aims at satisficing, rather than optimizing.

But when do we feel satisfied? In certain situations (as when foraging for food, or for sex) our body knows. In other situations, we have to find out. We may feel pleased if we pulled a lever which delivers one dollar, but not if we are told that the alternative would have delivered ten. In such a situation, we must learn what to aim for, whereas in the foraging case our germ line has done the learning already and the result is encoded in the genome. Natural selection operating in a population, or a learning rule based on individual trial and error, can cause an adaptation of the aspiration level.

It is easy to see how selection, or learning rules, lead to an optimal aspiration level when playing against nature. We are interested in exploring how adaptation works when playing against other players. In the repeated prisoner's dilemma game, for instance, a strategy called PAVLOV does very well (see Kelley *et al.* 1962; Colman 1995; Kraines & Kraines 1988; Nowak & Sigmund 1993). PAVLOV is a win–stay, lose–shift rule with an aspiration level lying somewhere between the two highest and the two lowest pay-offs. Is there any reason to assume that selection or learning will adapt the aspiration level precisely to this interval? How would such adaptive mechanisms fare in other games? We will assume that our players are 'blind robots' without any knowledge of the structure of the iterated game, except that they have two options. They need not even be aware of the existence of another player. Their only information is the pay-off which they obtain in each round.

In § 2, we shall briefly discuss some mechanisms for adapting the aspiration level, studying first the action of selection, and then two particularly simple learning rules, which are extremal cases of convex updating of the aspiration level, called YESTERDAY and FARAWAY. In § 3–5, we turn to the simplest games, symmetrical games between two players having two strategies each. We examine whether adaptive mechanisms lead to an efficient outcome for such 2×2 games. This is one aspect of a larger question, namely: when is satisficing optimizing?

Received 16 March 1999 *Accepted* 8 April 1999