to Victorina

Introduction

These lectures do not follow in the tradition of classical game theory. They do not employ concepts such as 'strategy', 'utility', 'payoff matrix', 'guaranteed result', and 'equilibrium point'. The reflexive game theory is intended to address problems different from those addressed by classical game theory. Its goal is to predict the individual choice made by a subject belonging to a group and outline the possibility for controlling this choice. We call this *reflexive control*. The term 'subject' refers to single individuals or to various types of organizations: political parties, military units, states, and even civilizations. Relations between the interests of a group and those of individual subjects are regulated by a rule called *the anti-selfishness principle*: a subject in a group while pursuing one's own goal may not cause harm to the group as a whole. This principle is as important in reflexive game theory as is the principle of guaranteed results in classical game theory.

The anti-selfishness principle renders actions unacceptable if they are advantageous to the subject but harmful for the group to which the subject belongs. It does not, however, forbid antisocial actions, if the subject does not undertake such actions in pursuing of his own interests. In this way, the subject's unselfishness may justify actions that cause harm to a group or society. Note that which actions are preferable for a subject and for a group are not determined in advance, but rather generated formally by the model itself.

An essential difference from classical game theory is that reflexive game theory makes special assumptions concerning the mental mechanism of choice. We assume that the subject possesses

a partially ordered set of self-images, that is, the subject has several images of the self, each of which may have images of the self, etc. The hierarchy of images is depicted by a special formula that we call *diagonal form*. Let us emphasize that the hierarchy of images is not arbitrary. It is always finite and predetermined by the graph of relationships among subjects. This is an important point. People have known for many centuries that the mental domain can be described by recursive chains of the type, "he knows that he knows that he knows . . . " Such chains have not been used in the models of human cognition, however, because there was no "stopping rule." That such a rule exists in reflexive game theory makes the theory viable.

Diagonal form defines a mental procedure of choice and, at the same time, the mathematical function describing that choice. Thus, having written a diagonal form according to empirical data, we automatically obtain the subject's choice function. One of the values in the diagonal form is interpreted as the subject's intention. Subsequently, we assume that the subject is intentional, which means that he has only such intentions as can become reality; thus, intentions are not assigned in advance. The intentional subject corresponds to an equation whose solution is interpreted as an alternative that can be chosen by the subject. The case where the equation has no solutions is interpreted as the subject's inability to make a decision under the given circumstances. The equation may have several solutions; each of them is considered to be a possible choice. Finally, there may be a case that any alternative can be the solution; we assume then that the subject has freedom of choice, i.e., the group does not impose constraints on his decision.

In this book, we provide examples applying reflexive game theory in the areas of personal relations, social life, politics,

international relationships, military decision making, and law.

Is it possible to test objectively this theory? The principle of falsification proposed by Karl Popper for the natural sciences can hardly be used in this case. General social-psychological theories cannot be considered wrong due to small numbers of failures, because their field of applications is defined only vaguely. The idea formulated by Donald Campbell (1997) fits better: in science, there is a principle of natural selection by which only those theories that are interesting for researchers survive. Reflexive game theory as well as classical game theory cannot be rejected due to numbers of failed predictions. Its fate depends on whether specialists use it and for how long.

The first step in constructing reflexive game theory was made more than forty years ago. The recursive chains, "I know that he knows that I know . . . " have been to found a model of the subject, with reference to the term "reflexive games" (Lefebvre, 1965, 1966, 1967). Then a special formal apparatus was developed for modeling human choice (Lefebvre, 1982), which applied the reflexive game theory to the analysis of specific situations (Lefebvre, 2001, 2007).

Important contributions to the theoretical understanding of reflexion have been made by numerous researchers. Tatiana Taran (1998, 2001) constructed a multivalued Boolean model for choosing social norms. Vladimir Krylov (2000) studied problems related to axiomatics of reflexive models. Yuli Schreider (1999) considered continuous-valued logics as languages of reflexion. Pavel Baranov and Vladimir Lepsky developed a formal model of the subject with reflexion and inner value (see Lefebvre, Baranov, Lepsky, 1969). Anatoly Trudolubov (1972) created a reflexive game model using dependency nets. Tim Kaiser and Stefan Schmidt (2008) found

relations between reflexive game theory and the theory of functors and categories. There were also two attempts to combine reflexive games with classical game theory. Novikov and Chkhartashvili (2003) included reflexive games into the formalism of classical game theory; Lefebvre (2001) included classical game theory within reflexive game theory. Future study will show whether such connections may be productive.

The book includes an Appendix. It contains problems and exercises relevant to the material described.

I am grateful to Viacheslav Filimonov, Tim Kaiser, Stefan Schmidt, Jonathan Farley, and Harold Baker for their advice and corrections. Invaluable help was given to me by Victorina Lefebvre, with whom I discussed my plans and the main ideas of the book. In addition, she prepared these lectures for publication and made many significant comments in the course of work. Without her participation this book would never have been completed.