Available online at www.sciencedirect.com
SCIENCE DIRECT®

Games and Economic Behavior 49 (2004) 49–80

GAMES and Economic Behavior

www.elsevier.com/locate/geb

Dynamic interactive epistemology

Oliver Board

Department of Economics, Amherst College, Amherst, MA 01002-5000, USA

Received 20 September 2002
Available online 31 January 2004

Abstract

The epistemic program in game theory uses formal models of interactive reasoning to provide foundations for various game-theoretic solution concepts. Much of this work is based around the (static) Aumann structure model of interactive epistemology, but more recently dynamic models of interactive reasoning have been developed, most notably by Stalnaker [Econ. Philos. 12 (1996) 133–163] and Battigalli and Siniscalchi [J. Econ. Theory 88 (1999) 188–230], and used to analyze rational play in extensive form games. But while the properties of Aumann structures are well understood, without a formal language in which belief and belief revision statements can be expressed, it is unclear exactly what are the properties of these dynamic models. Here we investigate this question by defining such a language. A semantics and syntax are presented, with soundness and completeness theorems linking the two.

JEL classification: C72; D82

Keywords: Interactive epistemology; Beliefs; Belief revision; Semantic; Syntactic; Language; Canonical structure

1. Introduction

It is well established both theoretically and empirically that strategic reasoning requires agents to form not just conjectures about each other's actions, but also about each other's knowledge and beliefs, which can then be used to infer what actions they might take. In particular, the implications of *common knowledge of rationality*, where all the agents are rational, all know they are all rational, all know that they know, and so on, have been extensively analyzed. More recently, epistemic foundations have been provided for game theoretic solution concepts such as Nash equilibrium (Aumann and Brandenburger, 1995).

E-mail address: ojboard@amherst.edu.

0899-8256/$ – see front matter
doi:10.1016/j.geb.2003.10.006

Comprehensive surveys of work in this area are provided by Dekel and Gul (1997) and Battigalli and Bonanno (1999).

Much of this work is based around the *Aumann structure* model (see Aumann, 1976), in which each agent's knowledge is represented by an information partition over a set of states, or possible worlds. For the purposes of the game theorist, however, Aumann structures have several important limitations. First, they describe a very strong concept of knowledge. An implication of modeling agents' epistemic states with information partitions is that everything they know is true, and that they have complete introspective access to this knowledge, i.e. they know everything they know (positive introspection), and they know everything they do not know (negative introspection). Negative introspection in particular has widely been considered inappropriate. More generally, it has been thought important to analyze agents' beliefs as well as their knowledge. And beliefs, unlike knowledge, can be false. These issues can be dealt with be replacing the information partitions with possibility correspondences (see e.g. Samet, 1990). Beliefs modeled by possibility correspondences at their most general do not satisfy any of the properties described above. By imposing certain restrictions on the correspondences we can recover these properties one by one.

The second problem with using Aumann structures to model rational play in games is that they are essentially static: the epistemic states that they model are fixed, while in dynamic games[1] agents have a chance to change their beliefs as the game progresses. In particular, conjectures about what strategies one's opponents might be playing can be revised as moves are observed. A stark illustration of the importance of such revisions is given by Reny (1993), who shows that once the possibility of belief change is taken into account, the game-theoretic wisdom that common knowledge of rationality implies backward induction in games of perfect information is undermined. As long as the information that an agent learns is consistent with what she already knew or believed, this problem can be handled in the existing framework. The agent's partition (or possibility correspondence) can be refined, in a manner analogous to Bayesian updating of probabilities, to take account of the new information. But, like Bayes rule, this process is not well defined when the information learned is incompatible with the agent's previous beliefs, i.e. she is *surprised*. And modeling the response to such surprises is crucial: to evaluate the rationality of strategies in a dynamic game, we must have a theory about what the players would believe at *every* node in the game, even though some of these nodes will typically be ruled out by the players on the basis of the information they possess at the beginning of the game.

Models of dynamic interactive reasoning have thus been developed. Stalnaker (1996) replaces the information partitions of the Aumann structure with *plausibility orderings* on the set of possible worlds, which encode information not just about each agent's current beliefs, but also about how these beliefs will be revised as new information is learned, even if this new information is a surprise (e.g. it takes the form of an unexpected move made by one's opponent). This seems to be a satisfactory resolution to the problem, and models

[1] That is, games in which there is a flow of information as the game proceeds. These games are commonly represented by the extensive form.