Journal of Mathematical Economics 51 (2014) 121–127

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Rational belief hierarchies☆

Elias Tsakas *

Department of Economics, Maastricht University, Netherlands

ARTICLE INFO

Article history:
Received 22 February 2013
Received in revised form
25 July 2013
Accepted 10 October 2013
Available online 21 October 2013

Keywords:
Belief hierarchies
Type spaces
Rational numbers

ABSTRACT

We consider agents who attach a rational probability to every Borel event. We call these Borel probability measures rational and introduce the notion of a rational belief hierarchy, where the first order beliefs are described by a rational measure over the fundamental space of uncertainty, the second order beliefs are described by a rational measure over the product of the fundamental space of uncertainty and the opponent's first order rational beliefs, and so on. Then, we derive the corresponding rational type space model, thus providing a Bayesian representation of rational belief hierarchies. Our main result shows that this type-based representation has the counterintuitive property that some rational types are associated with non-rational beliefs over the product of the fundamental space of uncertainty and the opponent's types, thus implying that the agent may attach an irrational probability to some Borel event even if she has a rational belief hierarchy.

1. Introduction

A belief hierarchy is a description of an agent's beliefs about some fundamental space of uncertainty, her beliefs about everybody else's beliefs, and so on. During the past few decades, belief hierarchies have become an integral tool of modern economic theory, often used to analyze games with incomplete information (Harsanyi, 1967–1968), as well as in order to provide epistemic characterizations for several solution concepts, such as rationalizability (Brandenburger and Dekel, 1987; Tan and Werlang, 1988), Nash equilibrium (Aumann and Brandenburger, 1995), and the correlated equilibrium (Aumann, 1987), just to mention a few.[1]

Belief hierarchies are in general very complex objects, consisting of infinite sequences of probability measures. This makes them in principle very hard to handle and sometimes even to describe, especially when it comes to high order beliefs. Having recognized this difficulty, Harsanyi (1967–1968) proposed an indirect Bayesian representation of belief hierarchies, known as the type space model. Formally, Harsanyi's model consists of a set of *types* for each agent and a continuous mapping from each type to the corresponding *conditional beliefs* over the product of the fundamental space of uncertainty and the opponent's type space. This structure induces a belief hierarchy for every type, thus reducing the infinite-dimensional regression of beliefs to a single-dimensional type. Mertens and Zamir (1985) and Brandenburger and Dekel (1993) completed the analysis by showing the existence of the universal type space, which represents all belief hierarchies satisfying some standard coherency properties.

In this paper we restrict attention to probabilistic beliefs that can take only rational values, e.g., we have in mind agents who do not hold beliefs of the form "*tomorrow it will rain with probability* $\sqrt{2}/2$". Such beliefs are modeled by Borel probability measures that attach a rational number to every Borel event. Throughout the paper, we call these probability measures *rational*.

Assuming that agents form rational beliefs over some underlying space of uncertainty Θ does not necessarily restrict the language they use in order to describe their beliefs, i.e., we remain within Harsanyi's framework which models the agents' language with the Borel σ-algebra of events in $\Delta(\Theta)$. This implies that infinite conjunctions/disjunctions are expressible, thus inducing a richer language than the ones typically used in logic.[2] As a consequence, our agents understand what it means to assign probability

☆ I am indebted to Itai Arieli, Christian Bach, Robin Cubitt, János Flesch, Amanda Friedenberg, Jayant Ganguli, Aviad Heifetz, Willemien Kets, Friederike Mengel, Ronald Peeters, Andrés Perea, Miklos Pinter, Arkadi Predtchednitski, Dries Vermeulen, Akira Yokotani, and the audiences in GAMES (Istanbul), LOFT (Sevilla), Nottingham (Economics) and Maastricht (EpiCenter) for useful comments and fruitful discussions on this paper. Financial support from the Marie Curie Fellowship (PIEF-GA-2009-237614) is gratefully acknowledged.

* Tel.: +31 681179174.
E-mail address: e.tsakas@maastrichtuniversity.nl.

1 For an overview of the epistemic game theory literature we refer to the textbook by Perea (2012) or the review article by Brandenburger (2008).

2 The standard syntactic models of logic typically assume that the language that describes the agents' beliefs is finitely generated by sentences of the form "*Eoccurs with probability at leastp*" where p is a rational number (Fagin and Halpern, 1994; Aumann, 1999). The latter induces an algebra of events in $\Delta(\Theta)$, which is obviously coarser than the Borel σ-algebra. Within this framework, Heifetz and Mongin (2001) provided a sound and complete axiomatization of Harsanyi's type-based models, while Zhou (2010) extended their analysis to the case of finitely additive type spaces.

0304-4068/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jmateco.2013.10.005

$\sqrt{2}/2$ to a Borel event $E \subseteq \Theta$, as the latter corresponds to the event $\{\mu \in \Delta(\Theta) : \mu(E) = \sqrt{2}/2\}$ which is Borel in $\Delta(\Theta)$. If on the other hand, the language used was finitely generated, similarly to the aforementioned models of logic, our agents would not be able to even understand what "*E occurs with probability* $\sqrt{2}/2$" meant, as such a sentence would not be expressible in the first place.[3] We further discuss the case of a finitely generated language later in the paper.

Even though our agents *can* express subjective beliefs that use non-rational probabilities, they refrain from actually doing so, as these beliefs are very complex. The idea is that agents are sophisticated enough to be able to understand every aspect of the environment, including what it means to form any kind of complex beliefs, but the latter is costly, and therefore they prefer to reason in simpler ways.[4] In fact, recent experimental findings on the understanding of non-rational numbers by students and mathematics teachers indicate that subjects find non-rational numbers very complex (Fischbein et al., 1995; Sirotic and Zazkis, 2007a,b). For instance, it is found that even though subjects know the concept and the structure of non-rational numbers, they still have a tendency to rely on decimal approximations which they find more intuitive. Thus, we find it plausible to assume that agents may use only rational numbers to express their subjective beliefs.

Supposing, as usual, that an agent thinks that everybody else reasons the same way as she does, it is not only her first order beliefs that are restricted but also her beliefs about everybody else's beliefs, and so on. For instance, besides Alexandra's beliefs not assigning probability $\sqrt{2}/2$ to E, she also does not put positive probability to Barney believing E with probability $\sqrt{5}/5$. In other words, her belief hierarchy is restricted to consist of a sequence of rational probability measures, where the first order rational beliefs are described by a rational measure over the underlying space of uncertainty, the second order rational beliefs are described by a rational measure over the product of the fundamental space of uncertainty and the opponent's space of first order rational beliefs, and so on. We call this infinite regression of probability measures a *rational belief hierarchy*.

Following Mertens and Zamir (1985) and Brandenburger and Dekel (1993), we construct a Harsanyi type space representation of rational belief hierarchies. However, as our main result (Theorem 1) shows, this Bayesian representation has an odd property. Namely, it contains rational types which are represented by *non*-rational probability measures over the product of the fundamental space of uncertainty and the opponent's rational type space. In other words, there is some Borel event in this product space to which this rational type attaches a non-rational probability even though every order of her belief hierarchy involves only rational probabilities. We find this result quite surprising, both from a technical as well as conceptual point of view.

The technical implication is rather straightforward. Namely, it says that, contrary to our intuition, rational belief hierarchies are not necessarily induced by types that are associated with a rational probability measure over the product of the underlying space of uncertainty and the opponents' types.

Regarding the conceptual contribution on the other hand, note that the only Borel events that receive a non-rational probability by a rational type correspond to sentences that describe the opponent's entire belief hierarchy, i.e., these events contain elements of the form "$\theta \in \Theta$ *occurs*", and "*the opponent's first order beliefs are* π_1", and "*the opponent's second order beliefs are* π_2", and so on. Thus, the previous result is relevant only for events that are expressible when the agent's language is modeled by the Borel σ-algebra, as it is the case in Harsanyi's model, but not when the agent has a finitely generated language like for instance in logic. But then, the natural question is whether we should actually care about this type of events. In other words, should we consider agents with a finitely generated language and finitely additive beliefs whose reasoning is exhausted with the formulation of their belief hierarchy, or should we also let agents have a countably generated language and countably additive beliefs who also form beliefs about the opponents' entire belief hierarchy? The answer to this question is far from being straightforward and has attracted the attention of several prominent researchers. For instance, while Savage (1972) postulates that finitely additive subjective beliefs should be used, Harsanyi (1967–1968) allows for countably additivity. The aim of this paper is not to provide a general answer to this question, but rather to point out that in the existence of a countably generated language agents may not be able to form subjective beliefs about every event in their language by using only rational numbers. Later in the paper, we also discuss our main result in the context of a finitely generated language, like the ones typically considered in logic.

The paper is structured as follows. In Section 2 we formally introduce the notion of rational probability measures and we prove some of their properties; Section 3 extends this framework to an interactive setting by introducing rational belief hierarchies. In Section 4 we construct a terminal rational type space model and prove our main result; Section 5 contains a concluding discussion.

2. Rational probability measures

We begin with some definitions and the basic notation. Let X be a Polish space, together with the Borel σ-algebra, $\mathcal{B}$.[5] As usual, $\Delta(X)$ denotes the space of probability measures on $(X, \mathcal{B})$, endowed with the topology of weak convergence.[6] For each $\mu \in \Delta(X)$, let $\text{supp}(\mu)$ denote the support, i.e., the smallest closed subset of X that receives probability 1 by μ.[7]

Consider the Borel probability measures that assign to every Borel event a rational number.

Definition 1. We define the set of *rational probability measures* by

$$\Delta^{\mathbb{Q}}(X) := \Big\{\mu \in \Delta(X) : \mu(B) \in \mathbb{Q},\ \forall B \in \mathcal{B}\Big\}. \quad (1)$$

We use rational probability measures to model an agent who does not hold beliefs of the form "*E occurs with probability* $\sqrt{2}/2$". Before moving forward, recall that the language that describes the

3 Restricting an agent's language resembles the structure typically considered in models of unawareness (Modica and Rustichini, 1999; Halpern, 2001; Heifetz et al., 2006). More specifically, in these models an agent is aware of a sentence if and only if she can express this sentence within the bounds of her language. Therefore, assuming that the agent's language cannot express non-rational probabilistic assessments is informally equivalent to the agent being unaware of the notion of non-rational numbers.

4 For instance, in a different framework, Eliaz (2003), Spiegler (2004) and Maenner (2008) study repeated games with players who prefer to form simple beliefs.

5 A topological space is called Polish whenever it is separable and completely metrizable. Examples of Polish spaces include countable sets endowed with the discrete topology and $\mathbb{R}^n$ together with the usual topology. Closed subsets of Polish spaces endowed with the relative topology are Polish. The countable product of Polish spaces, together with the product topology, is also Polish.

6 The topology of weak convergence, which is usually denoted by w^*, is the coarsest topology that makes the mapping $\mu \mapsto \int f d\mu$ continuous, for every bounded and continuous real-valued function, f. If X is Polish, then $\Delta(X)$ endowed with the topology of weak convergence is also Polish. For further properties of w^*, we refer to Aliprantis and Border (1994, Chapter 15).

7 If X is separable and metrizable, the support is unique (Parthasarathy, 1967, Theorem 2.1).