Iterated Expectations, Compact Spaces and Common Priors

Hellman, Ziv

7 June 2007

ITERATED EXPECTATIONS, COMPACT SPACES AND COMMON PRIORS

ZIV HELLMAN

zivyahel @ gmail.com

ABSTRACT: Extending to infinite state spaces that are compact metric spaces a result previously attained by Dov Samet solely in the context of finite state spaces, a necessary and sufficient condition for the existence of a common prior for several players is given in terms of the players' present beliefs only. A common prior exists iff for each random variable it is common knowledge that all its iterated expectations with respect to any permutation converge to the same value; this value is its expectation with respect to the common prior. It is further shown that the restriction to compact metric spaces is 'natural' when semantic type spaces are derived from syntactic models, and that compactness is a necessary condition. Many of the results are based on theorems from the general theory of Markov chains.

JEL Classification Numbers: C70, D82.

2000 Mathematics Subject Classification Numbers: 91A05, 91A06, 91A99, 60J27.

1. INTRODUCTION

The common prior assumption, ever since it was introduced into the study of games with incomplete information by Harsányi (1967-1968), posits that all women and men are 'created equal' with respect to probability assessments in the absence of information – hence the term common prior – and all differences in probabilities should, in principle, be traced to asymmetries in information received over time. The idea has become very pervasive, and in most applications of type spaces to economics it is assumed that players' beliefs can indeed be derived from a common prior by Bayesian updating.

A prior can be interpreted as the beliefs of a player in a previous period. In many models, however, any previous period is either fictional or irrelevant to the matter being studied. It is also clear that there are many plausible models of type spaces in which it is impossible for the players to have arrived at their current beliefs via updating from a common prior. This leads naturally to the question of whether a criterion can be identified by which one can tell, through the current beliefs of the players, that they have a common prior.

Aumann (1976), in his celebrated agreeing-to-disagree theorem, presented a necessary condition for the existence of a common prior in terms of present beliefs: if there is a common prior, then it is impossible to have common knowledge of difference in the beliefs of any given event. For many years, sufficiency remained an open question.

In the 1990s, several authors, by extending the notion of disagreement to differences in the expectation of a general random variable, were able to show that the impossibility of there being common knowledge of disagreement is also a sufficient condition for the existence of a common prior. Morris (1995) proved this result, in the context of a finite state space, by considering how the absence of common priors can affect the willingness to conduct trade in various trading environments. Feinberg (2000), utilising techniques closer to pure game theory, showed that both in the context of finite state spaces andof infinite but compact state spaces, a lack of common priors implies the existence of at least one common bet for which each player will subjectively assess that he or she has positive expectation – and showed that compactness is necessary for that result. Samet (1998b) proved the same result, with a finite state space, using separating hyperplane techniques. Finally, the extension of this result to compact spaces with a separating hyperplane argument – and appeal to the Riesz representation theorem – was attained in Heifetz (2001).

As Samet (1998a) pointed out, this criterion, based on disagreements, satisfactorily solves the question of how one can tell when players have a common prior, but it fails to express the common prior in a *meaningful* way; the fact that a disagreement cannot be common knowledge may guarantee the existence of a common prior, but it tells us nothing about this common prior. He then proceeded, in that same paper, to present a very different necessary and sufficient condition for the existence of a common prior that not only identifies the common prior when it exists, but also provides an epistemically meaningful interpretation to it.

This condition is expressed intuitively in Samet (1998a) in a colourful story. Imagine that Adam and Eve – who have both excelled in their studies at the same school of economics – are asked what return they expect on IBM stock. Having been exposed to different sources of information, we oughtn't be surprised if the two provide different answers. But we can then go on to ask Eve what she thinks Adam's answer was. Being a good Bayesian, she can compute the expectation of various answers and come up with Adam's expected answer. Likewise, Adam can provide us with what he expects was Eve's answer to that question. This process can continue, moving back and forth between Eve and Adam, theoretically forever. There are, in this example, two possible infinite sequences of alternating expectations, one that starts with Eve and one that starts with Adam.

Samet calls this process 'obtaining an iterated expectation', and shows that, when the relevant state space is finite, there exists a common prior if and only if both of these sequences converge to the same limit.

He achieves this result by representing Adam's beliefs[1] by a type matrix M_1 and Eve's beliefs by type matrix M_2. These then form two 'permutation matrices', $M_{\sigma_1} = M_2 M_1$, which is intended to be used for the process of obtaining iterated expectations starting with Adam, and $M_{\sigma_2} = M_1 M_2$, which does the same for the iterated expectations starting with Eve. It then turns out to be the case that both M_{σ_1} and M_{σ_2} are ergodic Markov matrices, and hence by standard results in Markov chain theory, each of them has a unique invariant probability measure, which may be labelled respectively p_1 and p_2. It is then shown that if $p_1 \neq p_2$, Adam and Eve cannot share a common prior. On the other hand, if $p_1 = p_2$, then not only is there a common prior, it has positively been identified – it is precisely $p := p_1 = p_2$.

This is a remarkable result, made all the more remarkable by the fact that it applies results developed in Markov theory for the study of stochastic processes to answer a question that seems not to be even remotely related. Samet (1998a), however, proves it only in the context of finite state spaces. Given the results in Feinberg (2000) and Heifetz (2001), which extend the other major criterion for a common prior to compact state spaces, it is natural to wonder whether the Samet characterisation can also be shown to hold in compact state spaces.

It is the goal of this paper to show that there is an affirmative answer to that question, and that, just as in Feinberg (2000), the compactness is necessary. The significance of such a result is clear, given that there are many models of interest which involve infinite state spaces and cannot be reduced to a finite space – we therefore extend the application of the Samet criterion to many models to which it previously could not be applied. The compactness limitation will also be shown not to be as constricting as it might appear at first glance, by appealing to an idea appearing in Feinberg (2000) in order to show that all 'natural' type-space models, in a sense that is made rigorous in the body of the paper, are compact.

[1] For the sake of simplicity here, we will make the mild technical assumption that the entire relevant state space is the meet of the Adam and Eve type space.

It should also be noted that the Samet criterion is significant because it provides, in principle, a way of calculating practically a common prior given a type space. In the finite state space context, one can form the type matrices and apply numerical solutions for calculating invariant probability measures in Markov chains – a subject of active research – in order to ascertain whether or not there is a common prior and if one exists, to identify it. Similarly, with the extension here of the Samet criterion to the more general compact spaces, it now becomes possible, given knowledge of the players' type spaces, in principle to estimate the expected values of random variables by use of numerical solutions, such as those appearing in e.g. Hernández-Lerma and Lasserre (2003).

The following rough correspondences exist between results in this paper and those that appear in Samet (1998a), save for the fact that the results in that paper are strictly limited to finite state spaces, whereas that restriction is lifted here:

Proposition 1 here is (roughly) an infinite state space version of Proposition 4 of Samet (1998a); Proposition 2 here corresponds to Proposition 5 of Samet (1998a); and similarly Proposition 3 corresponds to Proposition 2' and Proposition 4 to Theorem 1'.

OUTLINE OF THE PAPER

Definitions relating to type spaces and priors in general appear in section 2.1. Section 2.2 lists the major definitions and results from Markov chain theory that will be used in the sequel. The connection between Markov chain theory and type spaces is made explicit in section 3, which also introduces our definition of an 'admissible' type space. The main results of the paper are in sections 4 and 5. Section 6 relates to the idea of a 'natural' topology induced by syntactically modelling type spaces as opposed to semantic modelling, as discussed in Feinberg (2000), and shows that the results of this paper apply to all type space models induced from syntactic considerations. Section 7 shows that compactness is necessary for our results.

The paper concludes with some pointers to possible future avenues of research related to the subject matter covered here.

2. PRELIMINARY DEFINITIONS AND RESULTS

2.1 TYPE SPACES

A *type space* for a set of players is a tuple $\langle I, \Omega, \mathfrak{F}, (\kappa_i, t_i)_{i \in I} \rangle$. The set of players is denoted by $I = \{1,...,n\}$, where $n \geq 2$. Ω is a measurable space of arbitrary cardinality, whose elements are called *states*. The knowledge function κ_i on Ω is defined so that $\kappa_i(\omega)$ denotes player *i*'s knowledge when the true state of the world is ω. The symbol $\mathfrak{F}$ represents a σ-field of measurable events (subsets of Ω), and the beliefs – or probability measure – of each player at each state of the world is denoted by $t_i(\cdot \mid \omega)$.

The knowledge functions κ_i generate partitions Π_i of Ω, by

$$\Pi_i(\omega) := \{\omega' \mid \kappa_i(\omega') = \kappa_i(\omega)\}$$

The *meet* of $(\Pi_i)_{i \in I}$ is the partition Π of Ω which is the finest among all partitions that are coarser than Π_i for each *i*. For each $\omega \in \Omega$, $\Pi(\omega)$ denotes the element of the meet containing ω. A somewhat more constructive way to define the elements of the meet utilises the concept of 'reachability'. A state ω' is *reachable* from ω if there exists a sequence $\omega = \omega_0, \omega_1, \omega_2, \ldots, \omega_m = \omega'$ such that for each $k \in \{0,1,\ldots,m-1\}$, there exists a player i_k such that $\Pi_{i_k}(\omega_k) = \Pi_{i_k}(\omega_{k+1})$. It is well-known that $\omega' \in \Pi(\omega)$ iff ω' is reachable from ω, so that the relation of reachability can be used to define the partition Π.

Denote by K_i the σ-field generated by Π_i. This σ-field will be required to satisfy the property that for all players

$$K_i \subseteq \mathfrak{F}$$

so that the atoms of the knowledge partitions of each agent are $\mathfrak{F}$-measurable.

The probability distributions $t_i(\cdot \mid \omega)$ are required to satisfy three important properties:

i. $t_i(\cdot \mid \omega)$ is $\mathfrak{F}$-measurable, for each player and each state
ii. $t_i(\Pi_i(\omega) \mid \omega) = 1$, for each player and each state
iii. For every event $A \in \mathfrak{F}$ the function $t_i(A \mid \cdot)$ is K_i-measurable for each player

Property ii) can be described in words as saying that each player ascribes probability one to what he or she knows. Property iii) has the further important implication that each agent knows his or her own distribution, i.e. if the player has two different distributions in two states then he or she can distinguish between these states, so that for each $\omega' \in \Pi_i(\omega)$, $t_i(A \mid \omega') = t_i(A \mid \omega)$.

A *random variable* is a real-valued function on Ω. For a probability measure υ and a random variable *f* on Ω, the *expectation* of *f* with respect to υ is $\upsilon f := \int_\Omega f(\omega)\upsilon(d\omega)$.

For each player i and random variable f, i's expectation of f, denoted $E_i f$ is the random variable

$$(E_i f)(\omega) := \int_\Omega f(\varpi) dt_i(\varpi \mid \omega)$$

Given a type space, one can ask whether the space might have come to exist, in its current state, from a space with no information at all, by the players acquiring new information over time and updating their beliefs in a Bayesian manner. Each player's possible initial belief on the no-information primeval space is called a prior. In general, given player i's current type, there will not be a single prior from which the player could have arrived at the current state from the (hypothetical) primeval past – there will be a set of possible priors. A main question is then whether or not the agents have a common prior, meaning a possible initial identical belief that implies the differences in probability assessments currently seen amongst the players can be attributed solely to asymmetric information received over time.

More formally, an $\mathfrak{F}$-measurable probability distribution μ over Ω is a *prior* for player i if for every measurable event $A \in \mathfrak{F}$

$$\mu(A) = \int_\Omega t_i(A \mid \omega) d\mu(\omega)$$

In words, μ is a prior for player i if i's types $t_i(\omega)$ are the posteriors of μ conditional on i's information function t_i. A probability measure $P \subseteq \Delta(\Omega)$ is a common prior if it is a prior for each of the players $i \in I$.

As usual, χ_A is the indicator function of an event $A \in \mathfrak{F}$ taking the value 1 in A and 0 otherwise.

2.2 Markov Chains

A *stochastic kernel* or *transition probability function* on $(\Omega, \mathfrak{F})$ is a function P such that

i. $P(\cdot \mid \omega)$ is a probability measure for each fixed $\omega \in \Omega$
ii. $P(E \mid \cdot)$ is an $\mathfrak{F}$-measurable function on Ω for each fixed event $E \in \mathfrak{F}$

Given a transition probability function P, a *Markov chain* over $(\Omega, \mathfrak{F})$ is a discrete-time homogenous dynamical system that evolves in time in accordance with the n-step probability function $P^n(E \mid \omega)$ defined recursively by

$$P^n(E \mid \omega) = \int_\Omega P^{n-1}(E \mid \varpi) dP(\varpi \mid \omega) = \int_\Omega P(E \mid \varpi) dP^{n-1}(\varpi \mid \omega)$$

for all $E \in \mathfrak{F}$ and $\omega \in \Omega$.

Fix for the rest of this section a Markov chain M over $(\Omega, \mathfrak{F})$ with transition probability function P.

Let $\Delta(\Omega)$ denote the space of probability measures on Ω, with this space naturally outfitted with the induced weak* topology. It is possible to regard P as a function from $\Delta(\Omega)$ to $\Delta(\Omega)$, as follows. For each $\upsilon \in \Delta(\Omega)$, let

$$(\upsilon P)(E) := \int_{\Omega} P(E \mid \omega) d\upsilon(\omega)$$

Then P acts on $\Delta(\Omega)$ by way of $\upsilon \mapsto \upsilon P$. We will interchangeably write υM to stand for the same thing as υP, accepting the mild notational abuse of identifying M with its transition probability function. Using this notation, a probability measure υ is *invariant* with respect to the Markov chain M if $\upsilon = \upsilon M$. If such a measure exists, M is said to admit an invariant probability measure.

The transition probability function P can also be considered as operating on bounded functions in the following way. For each bounded integrable function f from Ω to **R**, let

$$(Pf)(\omega) := \int_{\Omega} f(\varpi) dP(\varpi \mid \omega)$$

Then Pf is a well-defined bounded function. Again, we will permit ourselves to write Mf to mean Pf.

If υ is an invariant probability measure with respect to the Markov chain M, then P can also be considered to be a linear operator on $L_1(\upsilon) := L_1(\Omega, \mathfrak{F}, \upsilon)$ into itself. We can then define, for any k and $f \in L_1(\upsilon)$

$$P^k f(\omega) := \int_{\Omega} dP^k(\varpi \mid \omega) f(\varpi)$$

We have in addition the concept of the Cesàro mean, defined as

$$P^{(n)} f(\omega) := \frac{1}{n} \sum_{k=0}^{n-1} P^k f(\omega)$$

If Ω has a topology τ, denote the class of bounded continuous functions with respect to τ from Ω to **R** by $C(\Omega)$. Then the chain M satisfies the *weak Feller* property if P maps $C(\Omega)$ to $C(\Omega)$.

Let φ be a non-trivial σ-finite measure for the space Ω. A Markov chain *M* is φ-*irreducible* if

$$\sum_{n=1}^{\infty} P^n(E \mid \omega) > 0$$

for all $\omega \in \Omega$ whenever $\varphi(E) > 0$ for $E \in \mathfrak{F}$.

We will make use of the following important theorems from the theory of Markov chains. These three theorems appear, in Hernández-Lerma and Lasserre (2003), respectively as Theorem 7.2.3, Proposition 4.2.2, and an amalgam of Theorem 2.3.4, Proposition 2.4.2 and Proposition 2.4.3:

THEOREM (Existence of invariant probability measure). Let Ω be a compact metric space, and let *M* be a Markov chain on Ω. Then *M* admits an invariant probability measure.

THEOREM (Uniqueness of invariant probability measure). Let *M* be a φ-irreducible Markov chain and suppose that *P* admits an invariant probability measure υ. Then υ is the unique invariant probability measure for *P*.

THEOREM (Birkhoff's ergodic theorem for Markov chains). Let *M* be a Markov chain and suppose that *P* admits an invariant probability measure υ. For every $f \in L_1(\upsilon)$ there is a function $f^* \in L_1(\upsilon)$ such that

$$P^{(n)} f \to f^* \quad \upsilon\text{-almost everywhere}$$

and

$$\int f^* \, d\upsilon = \int f \, d\upsilon$$

In addition, if υ is the unique invariant probability measure of *M*, then f^* is constant υ-almost everywhere, and $f^* = \int f \, d\upsilon$, υ-almost everywhere, so that

$$\text{the time-average } \lim_{n\to\infty} P^{(n)} f = \text{ the space-average } \int f \, d\upsilon, \ \upsilon\text{-a.e.}$$

3. TYPE SPACES WITHIN THE MARKOV FRAMEWORK

3.1 RELATING TYPE SPACES TO MARKOV CHAINS

In this section, we relate the concepts of type spaces and Markov chains.

First, note that by definition, the probability measure $t_i(\cdot|\cdot)$ of each player *i* satisfies the conditions for being a transition probability function, hence we can associate with each player a Markov chain based that probability measure, and label it M_i.

In general, given any two probability measures P_1 and P_2, one can further define a new probability measure $P_2P_1(E|\omega)$ by

$$P_2P_1(E|\omega) = \int_\Omega P_2(E|\varpi)dP_1(\varpi|\omega)$$

This obviously can be iterated any number of times. In particular, given a measure P, we recapitulate the definition of the infinite sequence of measures $\{P^n(\cdot|\omega)\}_{n\geq 1}$.

In our specific context, given any two players *i* and *j* and a measurable event *E*, the probability distribution $t_it_j(E|\omega)$ based on t_i and t_j is similarly defined by

$$t_it_j(E|\omega) = \int_\Omega t_i(E|\varpi)dt_j(\varpi|\omega)$$

In particular, given an element σ in *Sym*(*I*), the set of all permutations of the elements of *I*, define

$$t_\sigma := t_{\sigma(1)}\cdots t_{\sigma(n)}$$

iteratively, by using the above to define $t_{\sigma(n-1)}t_{\sigma(n)}$, then $t_{\sigma(n-2)}(t_{\sigma(n-1)}t_{\sigma(n)})$, and so forth. The Markov chain based on t_σ will be labelled M_σ.

We can now re-interpret various notions relating to a type space within the Markov framework.

First, note that for any function *f* on the state space, M_if is precisely the expectation of *f* in player *i*'s estimation. This is 'the primal case' (cf. Samet (1998a)), in the sense that the expectation is what is usually considered of economic significance and importance, as players choose their actions by comparing the relative expectations of functions.

Second, dual to this is, an invariant probability measure υ with respect to the Markov chain M_i is precisely a prior of player *i*. A common prior is a probability measure that is simultaneously invariant with respect to all $\{M_i\}_{i\in I}$.

A sequence $s = (i_1, i_2, ...)$ of elements of *I* is called an *I-sequence* if for each player *i*, $i = i_k$ for infinitely many *k*s. The *iterated expectation* of a random variable *f* with respect to the *I*-sequence *s* is the sequence of random variables $\{M_{i_k}...M_{i_1}f\}_{k=1}^{\infty}$.

Given the identification of $E_i f$ with $M_i f$, we can write, given a permutation σ of *I*,

$$M_\sigma := E_\sigma := t_\sigma := E_{\sigma(1)} \cdots E_{\sigma(n)} = M_{\sigma(1)} \cdots M_{\sigma(n)} = t_{\sigma(1)} \cdots t_{\sigma(n)}$$

The *iterated expectation of f with respect to* σ is defined by the sequence $\{E_\sigma^k f\}_{k=1}^\infty$. The iterated expectation of *f* with respect to σ is, of course, the iterated expectation of *f* with respect to the *I*-sequence

$$\sigma(1),...,\sigma(n),\sigma(1),...,\sigma(n),...$$

as defined above.

3.2 Admissible Type Spaces

A type space $\langle I, \Omega, \mathfrak{F}, (\kappa_i, t_i)_{i\in I} \rangle$ with a topology τ over Ω will be termed *admissible* if it satisfies the conditions:

i. Each state $\omega \in \Omega$ is contained in an event $A(\omega)$ such that $t_i(A(\omega) \mid \omega) > 0$ for all *i*.
ii. The correspondence $\omega \mapsto t_i(\cdot \mid \omega)$ is continuous with respect to the topology τ and the weak* topology of $\Delta(\Omega)$

These may seem to be artificial requirements, but they are 'natural' in a sense that will be explained in the section on syntactic models of type spaces.

Note that from previous definitions, it immediately follows that an admissible type space satisfies the properties that $t_i(A \mid \omega) > 0$ for every event *A* of non-zero measure and non-zero intersection with $\Pi_i(\omega)$, and that

$$\int_\Omega f(\varpi) dt_i(\varpi \mid \omega)$$

is continuous in ω for every $f \in C(\Omega)$.

If in addition to the above conditions (Ω, τ) is compact metric space, the type space $\langle I, \Omega, \mathfrak{F}, (\kappa_i, t_i)_{i\in I}, \tau \rangle$ will be called a *compact admissible space* in short. Nearly all the results in this paper will henceforth assume a compact admissible type space. For notational ease, $\langle I, \Omega, \mathfrak{F}, (\kappa_i, t_i)_{i\in I}, \tau \rangle$ will frequently be written simply as (Ω, τ).

4. COMMON PRIORS AND COMPACT ADMISSIBLE TYPE SPACES

Given any $Q \in \Pi$, the restriction of M_i to Q, for any player *i*, will be written as M_i^Q. Given a permutation σ in *Sym*(*I*), the restriction of M_σ to *Q* is similarly denoted by M_σ^Q.

LEMMA 1. Given a type space Ω satisfying property (i) of admissible type spaces, for any permutation σ of *I* and player *i*, for any arbitrary pair of states ω, $\varpi \in \Pi_{\sigma(i)}(\omega)$, $t_\sigma(A(\varpi) \mid \omega) > 0$.

PROOF. Let ω, $\varpi \in \Pi_{\sigma(i)}(\omega)$. By property (i) of admissible type spaces, there exist events $A(\omega)$ and $A(\varpi)$ such that, for $i < j \le n$, $t_{\sigma(j)}(A(\omega) \mid \omega) > 0$; for $1 \le k < i$, $t_{\sigma(k)}(A(\varpi) \mid \varpi) > 0$; and because $t_{\sigma(i)}(A(\varpi) \mid \varpi) > 0$ and $t_{\sigma(i)}(A(\varpi) \mid \omega) = t_{\sigma(i)}(A(\varpi) \mid \varpi)$ (as ω, $\varpi \in \Pi_{\sigma(i)}(\omega)$), it follows that $t_{\sigma(i)}(A(\varpi) \mid \omega) > 0$. Unravelling the recursive definition of $t_{\sigma(1)} \cdots t_{\sigma(n)}$, these facts taken together imply that $t_\sigma(A(\varpi) \mid \omega) > 0$.

PROPOSITION 1. For any permutation σ of *I*, and for any element *Q* of the meet of a compact admissible type space (Ω, τ), M_σ^Q has a unique invariant probability measure π_σ^Q.

PROOF. By the assumed properties of an admissible type space, the Markov matrix $M_{\sigma(i)}^Q$, for any *i*, satisfies the weak Feller property. The weak Feller property of the permutation matrix M_σ^Q, follows readily by its concatenation formation via $M_{\sigma(1)}^Q \cdots M_{\sigma(n)}^Q$. The compactness of the metric topology τ then guarantees the existence of at least one invariant probability measure for M_σ^Q, π_σ^Q, by application of a theorem quoted above.

Next, select an arbitrary event $E \subseteq Q$ such that $\pi_\sigma^Q(E) > 0$. By definition of a prior, we can readily select a state $\omega' \in E$ such that $t_{\sigma(1)}(E \mid \Pi_{\sigma(1)}(\omega')) > 0$ (otherwise there would be a contradiction to the assumption that $\pi_\sigma^Q(E) > 0$).

Let $\omega \in Q$ be selected arbitrarily. Due to the fact that ω and ω' share the same element of the meet, ω' is reachable from ω. This means that there exists a sequence $\{\omega = \omega_0, \omega_1, \omega_2, ..., \omega_m = \omega'\}$ such that for each $k \in \{0,1,...,m-1\}$, there exists a player i_k such that $\Pi_{i_k}(\omega_k) = \Pi_{i_k}(\omega_{k+1})$.

We can now define the following iterative process: by definition, there is a player i_0 such that $\Pi_{i_0}(\omega_0) = \Pi_{i_0}(\omega_1)$. At step 0 of the iterative process, we conclude from the lemma that $t_\sigma(A(\omega_1) \mid \omega_0) > 0$. At step 1, there is a player i_1 such that $\Pi_{i_1}(\omega_1) = \Pi_{i_1}(\omega_2)$, hence (again by applying the lemma) from $t_\sigma(A(\omega_2) \mid \omega_1) > 0$ and step 0, we arrive (from the definition of t_σ^2) at $t_\sigma^2(A(\omega_2) \mid \omega_0) > 0$.

At step j, there is a player i_j such that $\Pi_{i_j}(\omega_j) = \Pi_{i_j}(\omega_{j+1})$, hence from $t_\sigma(A(\omega_{j+1}) \mid \omega_j) > 0$ and step $j-1$, we arrive (from the definition of t_σ^j) at $t_\sigma^{j+1}(A(\omega_{j+1}) \mid \omega_0) > 0$.

At the end of the process, the conclusion is $t_\sigma^m(A(\omega_m) \mid \omega_0) > 0$. Finally, by a slight tweaking of the proof of lemma 1, from the fact that $t_{\sigma(1)}(E \mid \omega' = \omega_m) > 0$ and that for $k \in I$, $t_{\sigma(k)}(A(\omega_m) \mid \omega_m) > 0$, we can show that $t_\sigma(E \mid \omega') > 0$, so that $t_\sigma^{m+1}(E \mid \omega_0) > 0$. We thus conclude that M_σ^Q is π_σ^Q-irreducible, hence from the uniqueness of invariant probability measure theorem, π_σ^Q is unique. □

PROPOSITION 2. For a compact admissible type space Ω, the following conditions are equivalent.

i. π is a common prior on Ω
ii. π is an invariant probability measure of the Markov chain M_i for each $i \in I$
iii. π is an invariant probability measure of the Markov chain M_σ for each permutation σ

PROOF. This is the compact-space equivalent to Proposition 5 of Samet (1998a), and the proof is nearly identical.

Almost immediately from the definitions, i) and ii) are equivalent. That ii) implies iii) is quite easy – if $\pi t_i = \pi$ for each player, then one can successively calculate $\pi(t_{\sigma(1)} \cdots t_{\sigma(n)}) = \pi(t_{\sigma(2)} \cdots t_{\sigma(n)}) = \ldots = \pi t_{\sigma(n)} = \pi$ for any permutation σ.

It remains to show that iii) implies ii). Suppose iii) and let π be the invariant probability measure. Thus

$$\pi(t_1 t_2 \cdots t_n) = \pi$$

Multiplying from the right by t_1 gives

$$\pi(t_1 t_2 \cdots t_n t_1) = \pi t_1$$

So πt_1 is an invariant probability measure of $t_2 \cdots t_n t_1$. But by iii), π is an invariant probability measure of the Markov chain $M_2 \cdots M_n M_1$, and by the previous proposition $M_2 \cdots M_n M_1$ has a unique invariant probability measure. Thus, $\pi M_1 = \pi$ and similarly $\pi M_i = \pi$ for all i. ∎

COROLLARY. For each $Q \in \Pi$ there exists at most one common prior on Q. ∎

5. PERMUTATIONS, ITERATED EXPECTATIONS AND COMMON PRIORS

PROPOSITION 3. Given a compact admissible type space Ω, for each random variable f on Ω and permutation σ, $\lim_{n\to\infty} E_\sigma^{(n)} f$ exists, and on each element Q in Π is constant and is equal to $\pi_\sigma^Q f$, π_σ^Q-almost everywhere,.

PROOF. This follows from the previous propositions and Birkhoff's ergodic theorem, quoted above. ∎

PROPOSITION 4. Given a compact admissible type space Ω, with $\Pi = \{\Omega\}$, a common prior exists iff for each random variable f, the elements of $\{\lim_{n\to\infty} E_\sigma^{(n)} f \mid \sigma \in Sym(I)\}$ converge π_σ-almost everywhere to the same limit. Moreover, if π is the common prior, then this limit is πf, π-almost everywhere.

PROOF. As above, $\lim_{n\to\infty} E_\sigma^{(n)} f$ is constantly $\pi_\sigma f$, π_σ-almost everywhere, where π_σ is the unique invariant probability measure of M_σ on Ω. Thus, for each f, the limits for all σ are respectively π_σ-a.e. equal to each other iff for each f, $\pi_\sigma f$ are π_σ-a.e. constantly equal to one and the same value for all σ.

Clearly, if there is a probability measure π such that $\pi_\sigma = \pi$ for all σ, then $\pi_\sigma f$ are all equal to each other. In the other direction, if in particular for each $A \in \mathfrak{F}$, $\pi_\sigma \chi_A$ are all equal then there is a probability measure π such that $\pi_\sigma = \pi$ for all σ. This amounts, given previous claims, to saying that π is a common prior. ∎

We can summarise these results as follows:

THEOREM. Given a compact admissible type space Ω such that $\Pi = \{\Omega\}$, for each random variable f and permutation σ of the players, the iterated expectation of f with respect to σ converges and the value of its limit is common knowledge. Moreover, there exists a common prior if and only if for each random variable it is common knowledge that all its iterated expectations with respect to all permutations converge to the same value.

6. THE SYNTACTIC MODEL AND THE NATURAL TOPOLOGY

6.1 THE SYNTACTIC FOUNDATION OF TYPE SPACES

The syntactic construction of type spaces is developed and explored fully in Aumann (1999) and Feinberg (2000). Only a brief survey, sufficient for the needs of this paper, is presented here.

The syntactic construction begins with a list of letters in an alphabet. Each letter can be regarded as representing a natural occurrence, e.g. the letter x might represent the occurrence of 'the pound will rise against the dollar next month' or y might stand for 'the share price will fall at least by 3 and 1/4 next week'. The syntax is then generated by the letters and the players' beliefs and knowledge. It contains all possible formulae that correspond to statements about the occurrences, the knowledge of the players, their knowledge about the beliefs of others, and so forth.

More formally, a non-empty finite set of letters α constitutes an *alphabet*. A formula is a finite string of symbols obtained by applying the following four rules finitely often:

i. Every letter is a formula
ii. If f and g are formulae then so is $f \vee g$
iii. If f is a formula then so are $\neg(f)$ and $k_i(f)$ for all $i \in I$
iv. If f is a formula then so are $p_i^{\lambda}(f)$ for all $i \in I$ and every rational number $0 \le \lambda \le 1$

The intended semantic interpretation of these symbols is that $f \vee g$ stands for 'f or g', $\neg(f)$ stands for 'not f', $k_i(f)$ means 'player i knows f', and $p_i^{\lambda}(f)$ is 'agent i believes f with probability at least λ'.

A *closed* list L of formulae is a set of formulae which satisfy

$$(f \in L) \wedge (f \Rightarrow g \in L) \text{ implies } g \in L$$

A list L is *epistemically closed* if it satisfies

$$(f \in L) \text{ implies } k_i f \in L \text{ for all players } i$$

A *strongly closed* list is a list which is both closed and epistemically closed. Every formula that is in the strong closure of the list of formulae state in the appendix to this paper is a *tautology*.

A *state of the world* ω is a closed list that contains all tautologies and such that for every formula f it satisfies $\neg f \notin \omega$ iff $f \in \omega$. A state of the world is thus a full description of formulae that hold and are coherent in the sense of containing no contradictions.

A pairing of alphabet α and list of players *I* is called a *context*, and the set of all formulae generated by a context is called its *syntax*. The set of all states of the world defined in a syntax is denoted by Ω, and is called a *universal space*.

6.2 The Semantics of Syntax

Given a context and syntax, the set of states of the world can be assigned a natural semantic structure, so that we recapitulate the semantics of a type space. In order to do this, we need to define the full tuple $\langle I, \Omega, \mu, \mathfrak{F}, (\kappa_i, t_i)_{i\in I} \rangle$, to go along with the set of players *I* and the universal space with which we are given.

The knowledge functions κ_i assign to each state of the world ω the set of all formulae in ω that begin with the symbol k_i i.e. $\kappa_i(\omega) = \{f \in \omega \mid f = k_i g \text{ for some formula } g\}$. This is sufficient to generate the partitions by way of

$$\Pi_i(\omega) = \bigcap\nolimits_{f \in \kappa_i(\omega)} E_{k_i f}$$

For each formula *f*, E_f denotes the event that *f* obtains, meaning

$$E_f = \{\omega \in \Omega \mid f \in \omega\}$$

The σ-field generated by all the events E_f then forms $\mathfrak{F}$. The probability measures $t_i(\cdot \mid \Pi_i(\omega))$ are set by

$$t_i(E_f \mid \Pi_i(\omega)) = \sup\{\lambda \mid 0 \le \lambda \le 1, \lambda \in \mathbf{Q}, p_i^\lambda f \in \omega\}$$

and defining $t_i(\cdot \mid \Pi_i(\omega))$ in general as the extension of this to all of $\mathfrak{F}$.

6.3 The Natural Topology

Given a universal space Ω generated by a syntactic context, the sub-basis for the *natural topology* on Ω is the set of all events E_f where f is a formula.

With this basic definition to hand, we can prove a series of claims showing that a semantic type space generated by a syntactic context, with the natural topology, fulfils the assumptions needed for the main propositions proved in this paper. We hope thus to justify the supposition of those assumptions, in the sense that we can expect that 'most' of the type spaces of interest in applications will be generated from syntactic contexts.

CLAIM 1. Each state $\omega \in \Omega$ is contained in an event $A(\omega)$ such that $t_i(A(\omega) \mid \omega) > 0$ for all *i*.

PROOF. First, we show that for each player *i* there is a formula g_i and an event E_{g_i} such that $\omega \in E_{g_i}$ and $t_i(E_{g_i} \mid \omega) > 0$. To show this, assume that it is not true. Then, listing the denumerable formulae contained in ω as $\{f_1, f_2, ...\}$, $t_i(E_{f_j} \mid \omega) = 0$ for each *j*, so that $p_i^0 f_j \in \omega$ and $\neg p_i^\lambda f_j \in \omega$ for all $\lambda \in \mathbf{Q}$. But the tautologies (see the appendix) imply that $p_i^0 f_j \Rightarrow k_i p_i^0 f_j$ and $k_i p_i^0 f_j \Rightarrow p_i^1(p_i^0 f_j)$, which means that $p_i^1(p_i^0 f_j) \in \omega$, and we have reached a contradiction. Hence there is a formula g_i such that $t_i(E_{g_i} \mid \omega) > 0$ and $\omega \in E_{g_i}$.

Next, now that we have to hand the formulae $\{g_1, g_2, ..., g_n\}$, because $t_i(E_{g_i} \mid \omega) > 0$ we have a corresponding list $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ such that $\lambda_i > 0$ for all *i*, and $p_i^{\lambda_i} g_i \in \omega$. Starting from player 1 in the list, from the tautologies $g_1 \Rightarrow g_1 \vee g_2$, so $k_1(g_1 \Rightarrow g_1 \vee g_2)$, and hence $p_1^1(g_1 \Rightarrow g_1 \vee g_2)$, thus $p_1^{\lambda_1} g_1 \Rightarrow p_1^{\lambda_1}(g_1 \vee g_2)$. This means that $p_1^{\lambda_1}(g_1 \vee g_2) \in \omega$, hence $t_1(E_{g_1 \vee g_2} \mid \omega) > 0$. Similar reasoning leads to $t_2(E_{g_1 \vee g_2} \mid \omega) > 0$. Continuing this line of reasoning in an iterative manner, we can eventually set $h = g_1 \vee g_2 \vee ... \vee g_n$, and conclude $p_i^{\lambda_i}(h) \in \omega$, hence $t_i(E_h \mid \omega) > 0$ simultaneously for all *i*. Writing $A(\omega) := E_h$, the proof is concluded. □

CLAIM 2. The correspondence $\omega \mapsto t_i(\cdot \mid \omega)$ is continuous with respect to the topology τ and the weak* topology of $\Delta(\Omega)$.

PROOF. This is proved in the proof of Theorem 6 of Feinberg (2000). □

CLAIM 3. Ω is compact in the natural topology.

PROOF. This is proved in Lemma 1 of Feinberg (2000). □

CLAIM 4. The natural topology is Hausdorff.

PROOF. The sub-basis for the natural topology on Ω is the set of all events E_f where f is a formula. Hence each event E_f is an open set. Since for every two states there is a formula that holds in one state while its negation holds in the other state, the Hausdorff property follows immediately. □

CLAIM 5. The natural topology is regular.

PROOF. By standard results, a compact Hausdorff space is normal, which is a stronger property than regularity. □

CLAIM 6. The natural topology is second countable.

PROOF. This follows immediately from the fact that the sub-basis of the natural topology is countable, given the countable number of formulae. □

PROPOSITION 5. Every universal space Ω generated by a syntactic context, with the natural topology, is a compact admissible type space.

PROOF. Claims 1 and 2 establish that Ω satisfies the conditions of admissibility. Claim 3 shows the natural topology to be compact. Claims 4 through 6 together are the conditions for Urysohn's Metrization Theorem (see Munkres (1975)), so that we can conclude that Ω with the natural topology is a compact metric space. □

7. THE NECESSITY OF COMPACTNESS

In this section we demonstrate that the above results do not hold when the assumption of compactness is relaxed. This is accomplished by presenting a simple semantic model, with two players, in which consideration of a permutation does not lead to the existence of an invariant probability.

Consider two individuals, Anna and Ben, and a denumerable state space $\Omega = \{1,2,3,...\}$. Anna's partitions is $\{\{1\},\{2,3\},\{4,5\},...\}$ and Ben's partition is $\{\{1,2\},\{3,4\},\{5,6\},\ldots\}$. The meet in this case is all of Ω.

Ben's beliefs are always equal probabilities to the two states in each of his partition members. Anna's beliefs are also equal probabilities to the two states in her partition members, save for the probability 1 which is necessary for the one state partition member.

We can depict the beliefs of each of the two players in the form of infinite matrices:

$$Anna = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 & \ldots \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 & \ldots \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 & \ldots \\ . & . & . & . & . & 1/2 & \ldots \\ . & . & . & . & . & 1/2 & \ldots \\ . & . & . & . & . & 0 & \ldots \end{bmatrix}$$

$$Ben = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 \ldots \\ 1/2 & 1/2 & 0 & 0 & 0 \ldots \\ 0 & 0 & 1/2 & 1/2 & 0 \ldots \\ 0 & 0 & 1/2 & 1/2 & 0 \ldots \\ 0 & 0 & 0 & 0 & 1/2 \ldots \\ 0 & 0 & 0 & 0 & 1/2 \ldots \\ . & . & . & . & 0 \ldots \\ . & . & . & . & 0 \ldots \end{bmatrix}$$

Form the permutation matrix $M_\sigma := \text{Ben} \times \text{Anna}$

$$Ben \times Anna = \begin{bmatrix} 1/2 & 1/4 & 1/4 & 0 & 0 \ldots \\ 1/2 & 1/4 & 1/4 & 0 & 0 \ldots \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \ldots \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \ldots \\ 0 & 0 & 0 & 1/4 & 1/4 \ldots \\ 0 & 0 & 0 & 1/4 & 1/4 \ldots \\ . & . & . & 0 & 0 \ldots \\ . & . & . & 0 & 0 \ldots \end{bmatrix}$$

and note that it forms the following pattern: letting O stand for the set of positive odd integers, and regarding M_σ as a mapping on the domain $\mathbf{N} \times \mathbf{N}$, we start with $M_\sigma(1,1) = 1/2$, $M_\sigma(2,1) = 1/2$, and for each $j \in O$, $1/4 = M_\sigma(j, j+1) = M_\sigma(j+1, j+1) = M_\sigma(j+2, j+1) = M_\sigma(j+3, j+1) = M_\sigma(j, j+2) = M_\sigma(j+1, j+2) = M_\sigma(j+2, j+2) = M_\sigma(j+3, j+2)$. For all other values of k and l, $M_\sigma(k,l) = 0$.

Suppose now that there is an invariant probability distribution π with respect to M_σ. Let $\pi(1) = \alpha$. Then by the definition of invariant probability, it must also be the case that $\pi(2) = \alpha$, because $\pi(1) = 0.5(\pi(1) + \pi(2))$. Similar reasoning leads to the conclusion that $\pi(3) = \alpha$, $\pi(4) = \alpha$, … , $\pi(k) = \alpha$, …

Now, $\alpha \in [0,1]$, so that either $\sum_{k=1}^{\infty} \pi(k) = 0$ or $\sum_{k=1}^{\infty} \pi(k) = \infty$. In either case, π cannot be a normalised probability.

8. CONCLUSION

As stated in the introduction, in this paper we have extended most of the results of Samet (1998a) to compact metric spaces, shown that compactness is necessary for the results, and exhibited that never the less the results apply quite broadly to ‘nearly all’ models of interest by showing that every universal space generated by a syntactic context, with the natural topology, is a compact admissible type space.

It should be noted here that the results here do not extend all the results of Samet (1998a) to compact metric spaces. To be precise, the theorems of that paper, in the finite type space context, show that the existence of a common prior implies that for each random variable *f* it is common knowledge in each state that *all* the iterated expectations of *f*, with respect to all *I*-sequences *s*, converge to the same limit. The theorems of this paper show that, in the context of a compact admissible type space, the existence of a common prior implies that for each random variable *f* it is common knowledge in each state that the iterated expectations of *f with respect to each permutation* converge to the same limit. Whether our results on compact admissible type spaces apply to all iterated expectations with respect to all *I*-sequences remains an open question.

Finally, we note that Aviad Heifetz has expressed the following conjecture: first define an improper prior p_i for player *i* to be a σ-finite $\mathfrak{F}$-measurable measure over a type space Ω which satisfies the main condition of a prior, i.e. for every measurable event $A \in \mathfrak{F}$

$$p_i(A) = \int_{\Omega} t_i(A \mid \omega) dp_i(\omega)$$

Thus, an improper prior does not allow us to talk about the 'probabilities' that player i assigns to events at the ex-ante stage, but it still allows us to discuss the relative likelihood that he or she ascribes to pairs of events.

Then Heifetz has conjectured that a (not necessarily compact) type space admits a (possibly improper) common prior if and only if the Morris-Feinberg criterion holds. In analogy, one might similarly conjecture that a (not necessarily compact) type space admits a (possibly improper) common prior if and only if the Samet criterion holds. Whether or not either of these conjectures is true remains a subject for future research.

9. APPENDIX

LOGICAL TAUTOLOGIES

$$(f \vee f) \Rightarrow f$$
$$f \Rightarrow (f \vee g)$$
$$(f \vee g) \Rightarrow (g \vee f)$$
$$(f \Rightarrow g) \Rightarrow ((h \vee f) \Rightarrow (h \vee g))$$

KNOWLEDGE TAUTOLOGIES

$$k_i f \Rightarrow f$$
$$(k_i(f \Rightarrow g)) \Rightarrow ((k_i f) \Rightarrow (k_i g))$$
$$(k_i f) \Rightarrow (k_i k_i f)$$
$$(\neg k_i f) \Rightarrow (k_i \neg k_i f)$$

PROBABILITY TAUTOLOGIES

$$p_i^0 f$$
$$(k_i f) \Rightarrow (p_i^1 f)$$
$$(p_i^\lambda f) \Rightarrow (k_i p_i^\lambda f)$$
$$(p_i^\lambda f) \Rightarrow (p_i^\delta f) \qquad \lambda > \delta$$
$$p_i^\lambda (f \wedge g) \wedge p_i^\delta (f \wedge \neg g) \Rightarrow p_i^{\lambda+\delta} f \qquad \lambda + \delta \le 1$$
$$\neg p_i^\lambda (f \wedge g) \wedge p_i^\delta (f \wedge \neg g) \Rightarrow p_i^{\lambda+\delta} f \qquad \lambda + \delta \le 1$$
$$p_i^\lambda f \Rightarrow \neg p_i^\delta \neg f \qquad \lambda + \delta > 1$$
$$p_i^1 (f \Rightarrow g) \Rightarrow (p_i^\lambda f \Rightarrow p_i^\lambda g)$$

REFERENCES

Aumann, R (1976). “Agreeing to Disagree”, *Annals of Statistics* **4**, 1236–1239.

Aumann, R (1999). “Interactive epistemology II: Probability”, *International Journal of Game Theory* **28**, 301–314.

Feinberg, Y. (2000). “Characterizing Common Priors in the Form of Posteriors”, *Journal of Economic Theory* **91** (2), 127–179.

Harsanyi, J. C. (1967–1968). “Games with Incomplete Information Player by Bayesian Players, *I, II, III*,” *Management Science* **14**, 159–182, 320–334, 486–502.

Heifetz, A. (2001). “The Positive Foundation of the Common Prior Assumption”, Social Science Working Paper 1125, California Institute of Technology.

Hernández-Lerma, O. and Lasserre, J.B. (2003). *Markov Chains and Invariant Probabilities*, Basel : Birkhauser-Verlag.

Morris, S. (1994). “Trade with Heterogeneous Prior Beliefs and Asymmetric Information”, *Econometrica* **62** 1327–1347.

Munkres, J. (1975). *Topology*, New Jersey : Prentice-Hall.

Samet, D. (1998a). “Iterated Expectations and Common Priors”, *Games and Economic Behavior* **24**, 131–141.

Samet, D. (1998b). “Common Priors and the Separation of Convex Sets”, *Games and Economic Behavior* **24**, 173–175.