ECONOMIC RESEARCH REPORTS

LEARNING AND AGREEING TO DISAGREE WITHOUT COMMON PRIORS

BY

Yaw Nyarko

RR # 91-51 September, 1991

C. V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

Learning and Agreeing to Disagree Without Common Priors

Yaw Nyarko
New York University
Spring 1991

*I thank Professors Beth Allen and Ehud Kalai and members of the University of Pennsylvannia Economic Theory workshop for their comments and suggestions. I especially thank Professors Charles Wilson and Roy Radner for assistance and ecouragement in earlier iterations of this research. I am also very grateful to both the C.V. Starr Center and the Challenge Fund at New York University for their generosity.

Learning and Agreeing to Disagree Without Common Priors

Yaw Nyarko
New York University

Abstract.

We study the question of Common Knowledge and Agreeing-To-Disagree first discussed in the seminal paper of Aumann (1976) in a dynamic framework which is a generalization of Geanakoplos and Polemarcharkis (1982). We replace the Common Prior Assumption typically used in this literature with a generalization which merely requires that agents' priors satisfy an ex ante mutual absolute continuity condition. We obtain conditions for beliefs to converge. We show that beliefs and actions become common knowledge in the limit. We also discuss what happens to the Agreeing-To-Disagree results when priors are not common. In particular we show that limiting beliefs or actions or "announcements" equal that resulting from the (not necessarily common) prior conditional on common observations.

Mailing Address
New York University
Department of Economics
269 Mercer St.
New York, NY 10003.

II. Introduction. In this paper we study the question of Common Knowledge and Agreeing-To-Disagree first discussed in the seminal paper of Aumann (1976). Unlike most of the literature however, our emphasis will be on dynamic settings where we dispense with the Common Prior Assumption. Dynamic considerations arise naturally in the study of market economies over time or sequential games, where the question of whether there is "common knowledge" in the limit becomes important. In this paper we provide conditions for the convergence of beliefs and determine when those limiting beliefs are common knowledge and the when they are the same. In particular, we answer the question of what happens to the Agreeing-To-Disagree results when priors are not "Common."

We consider a dynamic setup without common priors which is a generalization of the model studied by Geanakoplos and Polemarcharkis (1982), where at each date "announcements" are made. Those announcements could for example be market prices in an economy where agents are incompletely informed about some "fundamentals." We show that the announcements converge and become common knowledge. We show that agents' limiting announcements will equal that resulting from their (not necessarily common) prior conditional on their "common observations."

In the model of Aumann (1976) it was shown that if agents have common priors and their posteriors are common knowledge then those posteriors must be the same. If we assume in our setup that the priors are common (i.e., the same) and the "announcements" are the posterior distributions we obtain original Aumann (1976) model, in a dynamic setup. Our result that the limiting announcements equal the prior conditional on the common observations then implies the Aumann (1976) Agreeing-To-Disagree result as a special case.

The common prior assumption begs the question of where the "common" priors come from. Imposing this assumption requires the imposition of a great deal of coordination of beliefs of agents, which runs against the very spirit of the incomplete information we seek to model. Furthermore, it has been shown in Nyarko (1991b) that when we index a game by the beliefs of each player-type, then the set of games which satisfy the common Prior assumption has much smaller dimensionality relative to the set of all such games. In addition, the relative dimension goes to zero as the number of players goes to infinity. Hence, "most" games violate the Common Prior Assumption.

We replace the Common Prior Assumption with a generalization which allows the priors to be different, but requires them to be mutually absolutely continuous so that there is ex ante agreement on which sets have probability zero. With this condition we may obtain the convergence of belief hierarchies over time. We believe that such a mutual absolute continuity condition is necessary[1] at the level of generality of this paper since without it one may easily obtain examples where beliefs cycle ad infinitum and hence do not converge. (See Nyarko, 1991a, for a simple example in the context of a single agent decision problem.)

Aumann (1976) studies common knowledge via finite partitions and "the meets" of those partitions. We define common knowledge in

Bayesian terms; i.e., an event is common knowledge if each agent assigns probability one to it; each agent assigns probability one to the event that other agents' beliefs assign probability one to the event; etc. Tan and Werlang (1988) show the equivalence of these two defintions under finiteness assumptions on the partitions. Nielson (1984) shows the equivalence for possibly infinite partitions via the use of completed Boolean Algebras and common priors. Brandenburger and Dekel (1987) show the equivalence via the use of proper conditional probabilities and posterior completed σ-algebras.

The proofs of the convergence of announcements available in the literature typically exploit critically the finiteness of the partitions. This is the case in Geanakoplos and Polemarcharkis (1982) and McKelvey and Page (1986). When the partitions of each agent is finite, then so is the join of those partitions. It is easy to see that the number of periods until the announcements no longer provide any new information is bounded above by the number of elements in the join. Given that the announcements have converged at some finite date, it then follows almost immediately that the announcements must be common knowledge from that date on. Nielson (1984) provides a proof of convergence of announcements when the partitions are not finite; that paper however imposes common priors.

Since we are concerned with dynamics we do not use meets of partitions or completions of information fields. In a dynamic setup the assumption of finiteness of partitions would be very restrictive since it implies that after finitely many periods there is no new information. We instead exploit the mutual absolute continuity properties in our generalization of the common prior assumption.

In our setup we identify a parameter, θ, which we refer to as the "fundamentals" which is the principal item over which agents have incomplete information. Agents have beliefs over θ, their first order beliefs. Since agents do not know the beliefs of other agents they will have beliefs about the first order beliefs of others; this will be their second order beliefs. Their third order beliefs will be their beliefs about the second order beliefs of other agents. Etc. Each agent will have a hierarachy of beliefs of all orders. This will be that agent's **type.**

The agents then receive observations or signals over time. These observations may for example be equilibrium prices which are a function of the actions (e.g., sales and purchases) made by the agents as a function of their beliefs. The observations provide information on the unknown parameter θ, and possibly also on the types of agents in the economy.

We use this setup with "types" to enable us to relate our work to the structure of Harsanyi (1968) and indeed to state our generalization of the Common Prior Assumption. Further by identifying such a parameter, we isolate that with respect to which we seek to state our limiting convergence and common knowledge results.

The paper is organized as follows. In section 2 some terminology in introduced. Section 3 constructs the types or hierarchies of beliefs. Section 4 defines a very general "observation process." Section 5 introduces the generalization of the Harsanyi Common Prior assumption we require. Section 6 states the result on convergence of belief hierarchies. All this is taken from Nyarko (1991c), which should be consulted for details and examples on the Generalized Harsanyi Condition and the convergence of beliefs result.

Section 7 introduces our definitions of common knowledge of events and random variables and proves some key results on common knowledge in the limit and relates common knowledge with measureability and "probability one events." Section 8 introduces the generalization of the model of Geanakoplos and Polemarcharkis (1982). Within the context of that model the question of Agreeing-to-Disagree without common priors is studied. All proofs are relegated to the appendix.

2. Terminology: I is the set of agents, assumed to be finite. Given any collection of sets $\{S_i\}_{i\in I}$, we define $S\equiv X_{i\in I}S_i$ and $S_{-i}\equiv X_{j\neq i}S_j$. Given any collection of functions $f_i:S_i \rightarrow Y_i$ for $i\in I$, $f_{-i}:S_{-i} \rightarrow Y_{-i}$ is defined by $f_{-i}(s_{-i}) \equiv X_{j\neq i}f_j(s_j)$. The Cartesian product of metric spaces will always be endowed with the product topology.

Given any metric space S we let $\mathcal{P}(S)$ denote the set of all probability measures over Borel subsets of S; all probability measures in this paper will be of this form. $\mathcal{P}(S)$ will always be endowed with the weak topology of measures. Given any μ in $\mathcal{P}(S)$, Supp μ denotes the support of μ, the smallest closed subset of S which has μ probability of one. μ(ds) denotes integration with respect to μ over S; in particular, if f is any real valued function on S, $\int f(s)\mu(ds)$ is the integral of f with respect to μ. If S=S'xS", $Marg_{S'}\ \mu$ is the marginal of μ on S'; and μ(ds') denotes integration with respect to the $Marg_{S'}\ \mu$.

3. The Hierarchy of Beliefs. We let Θ denote the "fundamental" parameter space, with some $\theta \in \Theta$ representing the "true" parameter value. Agents have incomplete information about θ. Each agent $i \in I$ will have beliefs about the value of θ; that will be that agent's first order belief. Since agents do not know the beliefs of others each agent will have beliefs about the first order beliefs of other agents; this will be that agent's second order beliefs. Inductively, an agent's k-th order belief will specify that agent's beliefs about the (k-1)th order beliefs of the other agents. An agent's **type** specifies the complete hierarchy of all orders of beliefs for that agent.

We now proceed to formally define the set T_i of types (or hierarchies of beliefs) for each agent i, (following Mertens and Zamir (1985)). In the construction below $T^k{}_i$ will denote the set of all k-th order beliefs of agent i, and T_i will denote the set of all hierarchies of beliefs all orders for the i-th agent. Those uninterested in the details of the formal construction may proceed to section 4.

The parameter space is assumed to be a complete and separable metric space. Construct the sets $\{T^k_i\}^\infty_{k=1}$ inductively as follows: $T^1_i \equiv \mathcal{P}(\Theta)$, and given T^k_i for $k \geq 1$, define $T_i^{k+1} \equiv \mathcal{P}(T^k_{-i} x \Theta)$. It should be clear that higher order beliefs of an agent should be related to the lower order beliefs of the **same** agent by some kind of projection operation; for example, if τ^1_i and τ^2_i are the first and second order beliefs of the same agent then τ^1_i should be the marginal distribution of τ^2_i on Θ. Indeed, the k-th order beliefs $\tau_i^k \in T_i^k$, determine a unique k-1 th order belief $\tau_i^{k-1} \in T_i^{k-1}$ via functions $\phi_i^k : T_i^{k+1} \rightarrow T_i^k$ defined inductively by setting for any subset B of Θ,

$$\phi^1_i(\tau^2_i)(B) \equiv \tau^2_i(\{T^1_{-i} x B\}) \text{ for all } \tau^2_i \epsilon T^2_i \tag{3.1}$$

(i.e., ϕ_i^1 is the operator that yeilds the marginal distribution on Θ from any joint distribution on $T^1_{-i} x \Theta$); and given ϕ_j^{k-1} for all $j \epsilon I$ define for any $\tau_i^{k+1} \in T_i^{k+1}$ and any subset B of $T_{-i}^{k-1} x \Theta$,

$$\phi^k_i(\tau_i^{k+1})(B) \equiv \tau_i^{k+1}(\{(\theta, \tau^k_{-i}) \mid (\theta, \phi_{-i}^{k-1}(\tau^k_{-i})) \in B\}) \tag{3.2}$$

The set of all possible types of agent i is then defined to be the set

$$T_i \equiv \{(\tau_i^1, \tau_i^2, \ldots) \epsilon X_{k=1}^{\infty} T_i^k \mid \tau_i^k = \phi_i^k(\tau_i^{k+1}) \text{ for all } k \geq 1\} \tag{3.3}$$

(One may be curious why we did not construct the types sets by $T_i^{k+1} = \mathcal{P}(T^k_{-i})$ as opposed to $\mathcal{P}(T^k_{-i} x \Theta)$. The reason is that we seek to allow the i-th agent to have beliefs under which various orders of beliefs of other agents are correlated. For example the i-th agent may think that if $\theta = \theta'$ then other agents have first order beliefs τ'^1_{-i} while if $\theta = \theta''$ then the other agents have first order beliefs

τ''^{1}_{-i}. The above construction allows for these correlations).

The set T_i can be shown to be homeomorphic to $\mathcal{P}(T_{-i}x\Theta)$ (see Mertens and Zamir (1985)). Hence each $\tau_i \in T_i$ may also be considered to be a probability measure on $T_{-i}x\Theta$. Indeed, since each agent is assumed to know its own type each $\tau_i \in T_i$ may also be considered a measure on $Tx\Theta$ where $T = X_{i\epsilon I}T_i$; such a measure τ_i on $Tx\Theta$ assigns probability one to agent i having beliefs τ_i. To recap., agent i's type τ_i has the following three equivalent definitions:

(a) A hierarchy of beliefs, i.e., a member of T_i in (3.3);

(b) a probability measure on $T_{-i}x\Theta$;

(c) a probability measure on $Tx\Theta$.

4. Dynamics and Learning.

4.i. The Observation Process. In the initial period, "date 0", each agent i will have some hierarchy of beliefs, $\tau_{i0} \in T_i$; there will also be some "true" value of the "fundamental" parameter, $\theta\epsilon\Theta$. The collection of belief hierachies of agents' and the true parameter vector will be denoted by $\Gamma = (\{\tau_{i0}\}_{i\epsilon I},\theta)\epsilon Tx\Theta$. This will represent the "true" state of the "economy" or "game" at date 0.

There is an observation process, $\{z_n\}^{\infty}_{n=1}$, which is a stochastic process taking values in some (complete and separable metric) space Z. The process (not necessarily i.i.d.) has a probability law or distribution P_Γ, which depends upon the vector $\Gamma = (\{\tau_{i0}\}_{i\epsilon I},\theta)$ $\epsilon Tx\Theta$.

The date n observation is a vector $z_n = \{z_{in}\}_{i\epsilon I}$. At the end of each date n agent i observes z_{in}. If all agents observe the same information then of course $z_{in} = z_{jn}$ for all i and j in I. Since each z_{in} itself may be a vector this formulation allows agents to observe some common signals (e.g., market prices) as well as private signals.

We shall suppose that the distribution of the observation process, P_Γ, is "common knowledge" (as a function of Γ) in a "constructive" sense. By this we mean that i uses P_Γ in forming beliefs about $z^\infty = (z_n)^\infty_{n=1}$ (i.e., P_Γ is i's beliefs about z^∞ conditional on $\Gamma \epsilon Tx\Theta$); also i believes that other agents use P_Γ in forming beliefs about z^∞; and i believes others believe others use P_Γ; etc.

The assumption that P_Γ is "common knowledge" is of course without loss of generality: One could always expand the definition of the parameter θ to include a specification of the probability law of the observation process, in which case P_Γ would necessarily be "common knowledge."

We assume that the sets Θ and Z are complete and separable metric spaces. This in turn implies that T and $Z^\infty = Z.Z.Z....$ are also complete and separable metric spaces if we endow sets of probability measures with the weak topology and endow product spaces with the product topology. (See Partharathy (1967, Chpt. II.6)).

The set of elements over which there is uncertainty is therefore the set

$$\Omega = Tx\Theta xZ^\infty \qquad (4.1)$$

Any $\omega \epsilon \Omega$ is a tuple $\omega = (\theta, \{\tau_{i0}\}_{i\epsilon I}, \{z_n\}^{\infty}_{n=1})$ which specifies the true parameter vector θ, the initial hierarchies of beliefs of the agents, $\{\tau_{i0}\}_{i\epsilon I}$, and the sample path of observations, $\{z_n\}^{\infty}_{n=1}$. $B(\Omega)$ will denote the set of Borel subsets of Ω. The evolution of the observation process is governed by the probability P_{Γ}. This shall be considered to be a probability distribution over Ω which assigns probability one to the true vector $\Gamma = (\{\tau_{i0}\}_{i\epsilon I}, \theta) \; \epsilon T x \Theta$ of types of agents and the parameter vector. P_{Γ} will be referred to as the **objective** probability distribution.

As explained in section 3, any initial hierarchy of beliefs, $\tau_{i0} \epsilon T_i$, for any agent i may be considered a probability measure on the set of $\Gamma \; \epsilon \; T x \Theta$. Since P_{Γ} is assumed known to each agent this results in a probability distribution over Ω. We denote this by $\mu_i(.|\tau_{i0})$; i.e., given $\tau_{i0} \; \epsilon \; T_i$, the associated measure over Ω is defined by

$$\mu_i(S|\tau_{i0}) \equiv \int_{T x \Theta} P_{\Gamma}(S) \tau_{i0}(d\Gamma) \text{ for each } S \; \epsilon \; B(\Omega) \qquad (4.2).$$

Let $\Im_{in} = \sigma(\{z_{i1}, \ldots, z_{in}\})$, the information (or σ-) field generated by agent i's observations $\{z_{i1}, \ldots, z_{in}\}$; let $\Im_{i\infty} = V^{\infty}_{n=1} \Im_{in}$, the information field generated by the entire sample path of agent i's observations. $\mu_i(.|\tau_{i0}, \Im_{in})$ is the probability $\mu_i(.|\tau_{i0})$ conditional on $\Im_{in}$, for n=1,2,.... and n=∞.

4.ii. Date n hierarchy of beliefs over Θ. It is easy to see how the date n conditional probability $\mu_i(.|\tau_{i0}, \Im_{in})$ results in a **date n** hierachy of beliefs $\tau_{in} \; \epsilon \; T_i$ for each agent i and date n=1,...,∞.

Indeed, agent i's first order beliefs at date n is simply Marg_θ $\mu_i(.|\tau_{i0},\mathfrak{I}_{in})$. If agent i could observe the types of others, τ_{-i0}, and their observations (in $\mathfrak{I}_{-in}$) agent i could compute those agents' first order beliefs at date n. Since $\mu_i(.|\tau_{i0},\mathfrak{I}_{in})$ is agent i's beliefs about other agents' types and observations, this induces beliefs about other agents' first order beliefs at date n; in particular this induces i's second order beliefs. It is easy to see how all higher beliefs may be constructed.

The formal construction is as follows: Fix any $\omega \in \Omega$. Let $\mu_i(.|\tau_{i0},\mathfrak{I}_{in})(\omega)$ be agent i's date n conditional probability at ω. Define $\tau^1{}_{in}(\omega) \equiv \mathrm{Marg}_\theta\ \mu_i(.|\tau_{i0},\mathfrak{I}_{in})(\omega)$; then $\tau^1{}_{in}(\omega)\in T_i{}^1$. Suppose we have constructed for each agent $j\in I$, that agent's k-th order belief $\tau^k{}_{jn}(\omega) \in T^k{}_j$ at each $\omega\in\Omega$ for some integer $k \geq 1$. Let $\theta(\omega)$ be the value of the "fundamental" parameter at $\omega\in\Omega$ (i.e., the projection of ω onto its θ-coordinate). Define $\tau_{in}{}^{k+1}(\omega) \in T_i{}^{k+1}$, by setting for each subset B of $T^k{}_{-i}x\Theta$, $\tau_{in}{}^{k+1}(\omega)(B) \equiv \mu_i(\{\omega'\in\Omega:(\tau^k{}_{-in}(\omega'),\theta(\omega')) \in B\}|\tau_{i0},\mathfrak{I}_{in})(\omega)$. The i-th agent's date n hierarchy of beliefs at ω is then defined by $\tau_{in}(\omega) \equiv \{\tau^k{}_{in}(\omega)\}^\infty_{k=1}$; it is easy to check that this lies in T_i.

τ_{in} is the date n beliefs of agent i about the other agents' date n hierarchy of beliefs over θ; i.e., i's beliefs about θ at date n; i's beliefs at date n about other agents' **date n** beliefs about θ; etc. Note in particular that τ_{in} is NOT the marginal of $\mu_i(.|\tau_{i0},\mathfrak{I}_{in})$ on T; the latter is belief hierarchy of agent i about other agents' **date 0** hierarchy of beliefs conditional on date n information, $\mathfrak{I}_{in}$.

5. The Generalized Harsanyi Condition. We now restrict the beliefs of agents to those satisfying a mutual absolute continuity property. Since the distribution of the observation process, P_Γ, is known to each agent (as a function Γ) it will suffice to place assumptions on agents' date 0 belief hierarchies. All probabilities in this section will therefore be with respect to date 0.

The "Harsanyi Doctrine" of Harsanyi (1968) is said to hold if each agent's type (and in particular, beliefs about other agents' types) is obtained by conditioning some joint probability (on the set of all types and parameter vectors) on it's own realized type. Under the Harsanyi Doctrine this joint probability is assumed to be the same for each agent. In the condition below we generalize the Harsanyi condition to allow the joint probabilities to differ across agents.

Given two probability measures q' and q" on a metric space S, we say that q' is absolutely continuous with repect to q", or q' << q" if for any (measureable) subset B of S, q"(B)=0 implies q'(B)=0; q' and q" are mutually absolutely continuous if q'<<q" and q"<<q'. Let $\overline{\Theta}$ and $\overline{T}_i$ be subsets of Θ and T_i.

(GH) For each $i \in I$ there exists a probability measure π_i on $T \times \Theta$, such that Supp $\pi_i = \overline{T} \times \overline{\Theta}$, $\tau_{i0} = \pi_i(.\,|\,\tau_{i0})$ for π_i-a.e. τ_{i0}, and for all i and j in I, π_i and π_j are mutually absolutely continuous.

Recall that τ_{i0} may be considered a probability measure on $T \times \Theta$; the same is true of $\pi_i(.|\tau_{i0})$. Condition (GH) requires the the equality of these two measures. The Harsanyi Doctrine would require that $\pi_i = \pi_j$ for all i and j, while the generalized Harsanyi condition (GH) allows these joint probabilities to be different. Condition (GH) further requires the joint probabilities, $\{\pi_i\}_{i \in I}$, to be mutually absolutely continuous with respect to each other.

6. Convergence of Beliefs Under the Generalized Harsanyi Condition. We now show that belief hierarchies converge under condition (GH). The convergence will hold on a set of sample paths with "probability one". Note however that the probability required is P_Γ, the objective probability and not any agent i's subjective probability $\mu_i(.|\tau_{i0})$.

The results of this section will **NOT** specify where beliefs converge to, and in particular the results here will not claim that there is convergence to complete information about the true parameter vector.

Let wlim denote the operation of taking the limit of a sequence of probability measures in the weak topology of measures. (See Billingsley, 1968, for more on this.) Recall that $\tau_{i\infty}$ is the hierarchy of beliefs under the limiting information field $\mathfrak{F}_{i\infty}$ for agent $i \in I$. Define for each $i \in I$ and each k and $n=1,\ldots,\infty$,

$$C^k_i \equiv \{\omega \in \Omega | \text{wlim}_{n\to\infty} \tau^k_{in} = \tau^k_{i\infty}\} \text{ and } C \equiv \bigcap_{i \in I} \bigcap^{\infty}_{k=1} C^k_i \qquad (6.1)$$

The set C^k_i is the set of sample paths where the i-th agent's k-th order beliefs converge; the set C is the set where all orders of beliefs of each agent converge.

We now show that $P_\Gamma(C) = 1$. This means that most sample paths (where "most" is measured with the objective probability P_Γ) lie in the set C. In particular, with P_Γ probability one there will be convergence of the belief hierarchies.

Theorem 6.1 (Convergence of Beliefs). Suppose condition (GH) holds. Let $\{\pi_i\}_{i\epsilon I}$ be the measures obtained under condition (GH) and let π be any measure mutually absolutely continuous with respect to the average of the agents' measures, $\Sigma_{i\epsilon I}\pi_i/(\#I)$, where #I denotes the cardinality of the set of agents, I. Then for π-almost every Γ ϵ $T\mathrm{x}\Theta$, $P_\Gamma(C) = 1$.

The above theorem states that for "most" values of the truth, Γ, the belief hierarchies of agents will converge on a set of sample paths with P_Γ-probability one, where P_Γ is the probability generated by "truth". The qualifier, 'for "most" values of the truth' means for a set of truths with π-probability one, where π is any measure absolutely continuous with respect to the average measure $\pi_{av} = \Sigma_{i\epsilon I}\pi_i/(\#I)$.

Now clearly we could set $\pi = \pi_{av}$ to obtain the required absolute continuity. Alternatively, since each π_i and π_j are mutually absolutely continuous (MAC) with respect to each other, each will be MAC with respect to the average measure $\pi_{av.}$ Hence we could set π equal to the π_i of any agent i.

7. Knowledge and Common Knowledge.

7.i. Some Definitions. From the definition of the sample space Ω, each $\omega\epsilon\Omega$ has associated with it (or generates) a type τ_{i0} and information or σ-field, $\mathfrak{F}_{in}$, for agent i at each date n=0,1,..., and n=∞. $\mu_i(.|\tau_{i0},\mathfrak{F}_{in})(\omega)$ represents the posterior probability of agent i at date n in the sample path $\omega\epsilon\Omega$. We shall use the terminology "at (ω,n)" to mean at date n in the sample path $\omega\epsilon\Omega$. For ease of exposition we shall let $\mu_{in}(.)(\omega)\equiv\mu_i(.|\tau_{i0},\mathfrak{F}_{in})(\omega)$. (Recall also that $\mathfrak{F}_{i0}$ is the trivial σ-field information representing no information).

In all of this section we shall let $X:\Omega\to S$ denote a generic (Borel measureable) random function taking values in some complete and separable metric space S. $1_{\{X(\omega)\epsilon A\}}$ denotes the indicator function, which equals one if $X(\omega)\epsilon A$ and equals zero otherwise; and for any $x\epsilon S$, $1_{\{X(\omega)=x\}}\equiv 1_{\{X(\omega)\epsilon\{x\}\}}$.

Definition 7.1 (common knowledge of a set). Agent i knows the set $D\ \epsilon\ B(\Omega)$ at (ω,n) if $\omega\epsilon D$ and $\mu_{in}(D)(\omega) = 1$. Define $K^1_{in}D \equiv \{\omega\epsilon\Omega|$i knows D at $(\omega,n)\}$ and $K^1_n \equiv \bigcap_{i\epsilon I}K^1_{in}$. Define inductively for r=1,2,...,

$$K_{in}^{r+1}D \equiv \{\omega\epsilon\Omega|\text{i knows } K_n^rD \text{ at } (\omega,n)\} \text{ and } K_n^{r+1} \equiv \bigcap_{i\epsilon I}K_{in}^{r+1}. \tag{7.1}$$

Also define $K^\infty_nD \equiv \bigcap^\infty_{r=1}K^r_nD$; if $\omega\epsilon K^\infty_nD$ the event or set D is said to be common knowledge at (ω,n).

Definition 7.2 (common knowledge of a random variable). Agent $i \in I$ knows the random function $X:\Omega \to S$ at (ω,n) if

$$\mu_{in}(\{\omega' \in \Omega \text{ s.t. } X(\omega')=x\})(\omega) = 1_{\{X(\omega)=x\}} \quad \text{for all } x \in S. \tag{7.2}$$

Define

$$K^1_{in}X \equiv \{\omega \in \Omega \mid i \text{ knows } X \text{ at } (\omega,n)\} \text{ and } K^1_n X \equiv \bigcap_{i \in I} K^1_{in}. \tag{7.3}$$

The random variable X is said to be common knowledge at (ω,n) if the set $K^1_n X$ is common knowledge at (ω,n).

Note that if we set $x=X(\omega)$ in eq. (7.2) then we obtain the conclusion that "i knows X at (ω,n)" implies that "i assigns probability one X taking the the value $X(\omega)$ at (ω,n)," and hence i knows the "true" value of X at (ω,n).

Definition 7.3. Let S be a separable metric space with metric d and let $S' \equiv \{s_1, s_2, \ldots\}$ be a separant. Let $B(s_m, 1/r)$ be the open ball of radius $1/r$ and center s_m; then define U(S) to be the set of all finite unions of the balls $B(s_m, 1/r)$ over all $m=1,2,\ldots$ and $r = 1,2,\ldots$. It is easy to check that U(S) is countable.

In the following lemma we show that knowledge of a random variable at (ω,n) is equivalent to knowing whether or not the random variable lies in any given set A in the class U(S) defined above.

Lemma 7.4. Agent i knows the random function $X:\Omega \rightarrow S$ at (ω,n) if and only if

$$\mu_{in}(\{\omega' \epsilon \Omega \text{ s.t } X(\omega') \epsilon A\})(\omega) = 1_{\{X(\omega) \epsilon A\}} \quad \text{for all sets } A \epsilon U(S). \tag{7.4}$$

Remark (Measureability). Fix any date n=1,2,... and any Borel measureable set D in Ω. Since the conditional probability $\mu_{in}(\cdot)(\omega)$ is by definition Borel measureable on Ω for each $i \epsilon I$, the set $K_n^1 D$ in definition 7.1 where each agent knows the set D, can be shown to be Borel measureable. This in turn will imply that the sets $K_n^r D$ of r-th level knowledge of D are each Borel measureable for each r= 1,2,...,∞. For any set $A \epsilon U(S)$ of definition 7.3, define $W_A \equiv \{\omega \epsilon \Omega \mid (7.4)$ holds for the given set A$\}$. Then since μ_{in} is Borel measureable, so is the set W_A. Since the class of sets in U(S) is countable, this implies that $W \equiv \bigcap_{A \epsilon U(S)} W_A$ is also a Borel measureable set. Lemma 7.4 then implies that the set $K^1{}_{in}X$ where i knows the random function X is Borel measureable. The Borel measureablility of the sets in definition 7.1 in turn implies that the sets where there is higher level knowledge of X are all Borel measureable.

7.ii. Common Knowledge Under condition (GH). In the next Proposition and the rest of this subsection we use the Generalized Harsanyi condition (GH). Recall that under that condition, before agent i observes its type τ_{i0}, that agent has a prior over the set of types and parameter vectors given by $\pi_i \epsilon \mathcal{P}(Tx\Theta)$. This in turn induces a prior, μ_i, over the set of sample paths Ω defined by

$$\mu_i(B) \equiv \int P_\Gamma(B)\pi_i(d\Gamma) \text{ for each } B \in B(\Omega). \tag{7.5}$$

Let π be any measure as in Theorem 6.1 (i.e., $\pi \in \mathcal{P}(T\times\Theta)$ and is any measure mutually absolutely continuous with respect to $\pi_{av} \equiv \Sigma_{i\in I}\pi_i/(\#I)$). We let μ_π be the probability over Ω induced by π and P_Γ; i.e., we define

$$\mu_\pi(B) \equiv \int P_\Gamma(B)\pi(d\Gamma) \text{ for each } B \in B(\Omega). \tag{7.6}$$

If a subset B of Ω has μ_π-probability one then for π-a.e. Γ, $P_\Gamma(B)=1$. Under the mutual absolute continuity condition of assumption (GH), (which we refer to as the M.A.C. condition) each μ_i and μ_j is mutually absolutely continuous with respect to each other and with respect to μ_π.

Under condition (GH) each agent's set of all "conceivable states of the world", ω, is a subset of the support of μ_π. The next proposition states that if in every such "conceivable state of the world" each agent knows some set D then that set must be common knowledge at each such state.

Proposition 7.5 (Knowledge a.e. of a set implies Common Knowledge.) Fix an $n=0,1,2,\ldots,\infty$. Suppose that for μ_π-a.e. $\omega\in\Omega$ each agent $i\in I$ knows the subset D of Ω at (ω,n). Then for μ_π-a.e. ω, the set D is common knowledge at (ω,n).

An immediate implication of the above lemma is the following:

Corollary 7.6. (Knowledge a.e. of X implies X is common Knowledge). Suppose that for μ_π-a.e. ω and some $n=0,1,2,\ldots,\infty$, each agent i knows the random function $X:\Omega\to S$ at (ω,n). Then for μ_π-a.e. ω, X is common knowledge at (ω,n).

Suppose that the random variable $X:\Omega\to S$ is measureable with respect to $\mathfrak{I}_{in}$ for all $i\in I$ at some $n=0,1,\ldots,\infty$. This would be the case if each agent observed the random variable X at date n in every state of the world. Is X common knowledge at (ω,n)? Due to "probability zero subtleties", the answer to this question is no! The measureability of a random variable is determined solely by the σ-field, $\mathfrak{I}_{in}$ in our case, and is a concept independent of the underlying probability measure, μ_{in} in our case. Knowledge and common knowledge use the conditional probabilities in their definition. However conditional probabilities are defined only up to "probability zero." On "probability zero" sets the conditional probabilities may be arbitrarily defined and hence can do "strange things."

In particular, on a set of sample paths, ω, with μ_i probability zero the conditional expectation at ω, $\mu_i(.|\tau_{i0},\mathfrak{I}_{in})(\omega)$ need not assign probability one to X taking the observed value $X(\omega)$ at ω! On such sample paths agent i has beliefs which contradict observations. In the language of Blackwell and Dubins (1975) such probability measures are not "proper."

To avoid this problem we may proceed in one of two ways. On the one hand we may assume that agents have proper beliefs and have proper beliefs of other agents having proper beliefs, etc. We shall not follow this route, as this may itself lead to other sutle complications (see Brandenburger and Dekel, 1987). Instead we shall impose condition (GH). With that we obtain the following:

Proposition 7.7. (Common Observations are Common Knowledge.) Suppose that for some $n=0,1,2,\ldots,\infty$, the random function $X:\Omega\to S$ (taking values in the complete and separable metric space S) is measureable with respect to $\mathfrak{F}_{in}$ for each $i\in I$. Then for μ_π-a.e. ω, X is common knowledge at (ω,n).

The following is a corollary of the above proposition. It states that if a sequence of random variables is observed at each date by each agent and if the random variable converges in "every conceivable state, ω," then the limiting random variable is common knowledge.

Corollary 7.8. (Limit of Common Observations are Common Knowledge). Let $\{X_n\}^{\infty}_{n=1}$ be a sequence of random functions on Ω taking values some (complete and separable) metric space S. Suppose for each $n=1,2,\ldots,$ and $i\in I$ X_n is $\mathfrak{F}_{in}$-measureable. Suppose further that with μ_π-probability one $\lim_{n\to\infty}X_n$ exists and equals some limiting random function X_∞. Then for μ_π-a.e. $\omega\in\Omega$, the limiting random variable X_∞ is common knowledge at (ω,∞) (i.e., under the limiting σ-fields $\mathfrak{F}_{i\infty}$).

8. The Geanakoplos-Polemarcharkis "Oral" Exchange.

8.i. The Model. In the model of Geanakoplos and Polemarcharkis (1982) (henceforth the G-P model) there is a finite collection, I, of agents each with incomplete information about some parameter θ in some parameter set Θ. Let $\{\tau_{i0}\}_{i\epsilon I}$ denote the initial (date 0) hierarchy of beliefs of the agents over θ. Each agent $i\epsilon I$ receives some private information, z_{i1}, at date one about θ. Agents use this information to revise their belief hierarchies over the parameter θ. (We will describe formally the updating process later). In the next period each agent announces to all others their posterior probability of the parameter lying in some fixed set E.

Since agent j's announcement will provide agent i information about the private information observed by j, agent j's announcement will be used by i to update i's beliefs about θ. In the next period, after each agent has revised its beliefs using the previous period reports, each will report their new reports to all agents. This again provides information about θ to each agent. Belief hierarchies will again be updated, and new announcements will be made based on the updated beliefs. This process is continued in each period.

We shall generalize the original G-P model in many directions. First, we shall suppose that at the end of each date n, if agent i has beginning of period n hierarchy of beliefs $\tau_{in-1} \epsilon T_i$, the announcement is some function of vector of date n hierarchy of beliefs of agents, $\delta(\tau_{n-1})$, where $\delta:T \rightarrow A$ maps the date n vector of types of agents into the set of possible date n announcements, A, a comlete and separable metric space. In many cases the announcements are a vector, $\delta=\{\delta_i\}_{i\in I}$, where $\delta_i:T_i\rightarrow A_i$ is i's announcement as a function of its type. Some typical types of announcements that have been studied in the literature are as follows: Either,

(a) each agent i announces the posterior probability it assigns to θ lying in some fixed set E; i.e., $\delta_i(\tau_{iN}) = \tau^1_{iN}(E)$ and $\delta=\{\delta_i\}_{i\in I}$. This is of course the announcements that are used in the original G-P model as described above. Or,

(b) each agent i announces it's posterior distribution on θ; i.e., $\delta_i(\tau_{iN}) = \tau^1_{iN}$ where τ^1_{iN} is the first order belief (or first coordinate) of τ_{iN}. Or,

(c) each agent i announces it's posterior expectation of θ; i.e., $\delta_i(\tau_{iN}) = \int\theta d\tau^1_{iN}$, where here θ is assumed to lie in some compact subset of a finite dimensional Euclidean space. Or,

(d) each agent observes the average of the announcements functions $\delta(\tau_n)\equiv\Sigma_{i\in I}\delta_i(\tau_{in})/(\#I)$ of the values $\delta_i(\tau_{in})$ where either δ_i is as in (a) or is as in (c) above (with θ real-valued). This is typically the case in models with economic markets where agents observe prices which are aggregates of agents' beginning-of-period expectations (see, e.g., McKelvey and Page, 1986).

As a second generalization of the original G-P model, we allow agents to receive at date n not only the vector of date n announcements of all the agents, $\delta(\tau_{n-1})$, but possibly some other supplementary private information, z'_{in}, which may be correlated with the true parameter vector. Hence, at the end of date n the observation of agent i is the vector $z_{in} = (\delta(\tau_{n-1}), z'_{in})$.

Further, unlike the original G-P model we shall **NOT** suppose that agents begin with common priors. We shall also **NOT** assume that the private information that agents observe can take only finitely many values (and in particular, using the language of Geanakoplos and Polemarcharkis, 1982, we shall not assume that agents' information partitions are finite).

8.ii. The Formal Construction. We now show formally how the model described in previous sub-section can be made to fit the precise framework of section 4. Those uninterested in the formal details may proceed to section 8.iii. At date 0 there will be some true vector of agent types and parameter vector $\Gamma=(\theta,\tau_0) \in \Theta\times T$. The date one observations are the vector $z_1 \equiv (z_{i1})_{i\in I}$ taking values in some complete and separable metric space Z and governed by some probability distribution which is some function of θ; i.e.,

$$z_1 \sim P_1(.|\theta) \tag{8.1}$$

z_{i1} denotes the coordinates of z_1 observed by agent i.

We proceed inductively. Suppose for some N≥1, we have defined the joint probability distribution of the observations for the first N periods, $z^N \equiv (z_1,\ldots,z_N) \in X^N_{n=1}Z$, as a function of the date 0 types and the parameter vector $\Gamma \in \Theta\times T$:

$$z^N \sim P^N(.|\Gamma). \tag{8.2}$$

First we show how (8.2) induces a beginning of date N+1 hierarchy of beliefs, τ_{iN}, for each agent i: Recall that τ_{i0} may be considered a probability measure on $Tx\Theta$. Define τ^*_{iN} to be the probability distribution induced on $Tx\Theta xX^N_{n=1}Z$ by τ_{i0} and $P^N(.|\Gamma)$; i.e., given any (Borel measureable) subset B of $Tx\Theta xX^N_{n=1}Z^n$

$$\tau^*_{iN}(B) \equiv \int P^N(B|\Gamma)\tau_{i0}(d\Gamma) \tag{8.3}$$

Let $\tau^*_{iN}(.|z_i^N)$ denote the probability τ^*_{iN} conditional on i's observations $z_i^N \equiv (z_{i1},\ldots,z_{iN})$. One then proceeds just as in the construction in section 4.ii. (with $\tau^*_{iN}(.|z_i^N)$ taking the place of $\mu_i(.|\tau_{i0},\mathfrak{F}_{in})$). This will result in date N+1 infinite hierarcy of beliefs, $\tau_{iN} = g_i^N(z_i^N,\tau_{i0})$, for each agent $i\epsilon I$ which is some function, g_i^N, of that agents' date 0 hierarchy of beliefs, $\tau_{i0}\epsilon T_i$, vector of observations, z_i^N, from date 1 to N.

The date N+1 observation is then the tuple $z_{N+1} = (\delta(\tau_N),z'_{N+1}) = (\delta(g^N(z^N,\tau_0)),z'_{N+1})$ where $\delta(\tau_N)$ is the date N vector of announcements of agents and z'_{N+1} is the vector of other supplementary private information the agents may receive. This defines the probability distribution, $P_{N+1}(.|z^N,\Gamma)$, of z_{N+1}. The joint distribution over Z^{N+1} is then defined by $P^{N+1}(dz^{N+1}|\Gamma) \equiv P_{N+1}(dz^{N+1}|z^N,\Gamma).P^N(dz^N|\Gamma)$. By induction we have this for all N=1,2,...

The sequence of joint distributions on finite dimensional subsets of Z^∞ (the cylinder sets) extends naturually to a unique probability, $P_\Gamma(.)$, over the entire sample space Z^∞. (One may consult, e.g., Laha and Rohatgi (1979) for more on unique extentions of measures).

8.iii. The Convergence of Announcements and Common Knowledge In the Limit. In all of this sub-section we assume condition (GH) holds. We know from Theorem 6.1 that the belief hierarchies of the agents, $\{\tau_{iN}\}_{i\in I}$, converge over time. The announcements of agents are some function, δ, of those hierarchies of beliefs. Hence if those functions are continuous we obtain that the announcements also converge over time. In examples (b) and (c) of the previous sub-section it is easy to see that the required continuity holds. Under example (a), we are not assured of the required continuity. However even for that example an easy direct martingale argument shows that the announcements converge. In particular we have the following:

Proposition 8.1. (Convergence of Announcements): Suppose condition (GH) holds in the model of section 8.1, and let π be any measure on $T\times\Theta$ as in Theorem 6.1. Suppose further that the announcement function, δ, is either of the form of one of examples (a)-(d) of the previous section, or is continuous. Then for π-a.e. $\Gamma\epsilon T\times\Theta$, $\lim_{n\to\infty} \delta(\tau_n) = \delta(\tau_\infty)$ with P_Γ-probability one.

At the beginning of date n agent i will have a type $\tau_{in-1} \epsilon T_i$. The announcement $\delta(\tau_{n-1})$ is then made for all to hear. At the end of date n, each agent will have information σ-field $\mathfrak{F}_{in}$. Hence for each n>1, the date n announcement, $\delta(\tau_{n-1})$, is measureable with respect to $\mathfrak{F}_{in}$ for each i. We know from Proposition 8.1 that $\delta(\tau_n)$ converges to $\delta(\tau_\infty)$ with μ_τ-probability one. From Corollary 7.8 this in turn implies that for μ_τ-a.e. ω, $\delta(\tau_\infty)$ is common knowledge at (ω,∞). Hence we have:

Proposition 8.2. (Common Knowledge of Limit announcements). Suppose that the announcement function, δ, is continuous or is of the form of example (a)-(d) in the section 8.i. Then for μ_τ-a.e. ω, the limiting announcement, $\delta(\tau_\infty)$, is common knowledge at (ω,∞).

Let us suppose for a while that agents announce their posterior (first order) beliefs on θ at each date as in example (b) of section 8.i. Proposition 8.2 implies that the limiting belief hierarchies, $\{\tau^1_\infty\}_{i\in I}$, will be common knowledge at (ω,∞) for almost every $\omega\in\Omega$. But what exactly does common knowledge of τ^1_∞ imply? First of course i knows each agents beliefs about θ. This defines i's second order beliefs. Each agent knows that i knows this (since τ^1_i is common knowledge). In particular, each agent knows i's second order beliefs. This of course defines each agents third order beliefs. Proceeding inductively this way we see that each agent will know the entire belief hierarchy of each agent. Further each agent will know that others know the entire belief hierarchy; etc. In particular we have the following (where by common knowledge below we of course mean in the sense of definition 7.2).

Corollary 8.3. (Common Knowledge of Limit Belief Hierarchies). Suppose that announcements are as in example (b) of section 8.i. Then for μ_τ-a.e. ω, the vector of belief hierarchies of agents, $(\tau_{i\infty})_{i\epsilon I}$, is common knowledge at (ω,∞).

We stress at this stage that common knowledge of the belief hierarchies does not imply that they are the same. Indeed at date 0, each agent $i\epsilon I$ may be "born" with some hierarchy of beliefs, τ_{i0}. These may be common knowledge at date 0. Suppose that the information signals at date 1, $\{z_{i1}\}_{i\epsilon I}$, contain no information on θ (e.g., they are constant numbers); suppose also that there are no supplementary private information over time. Then the process of announcing posterior distributions at each date conveys no new information about θ. Agents never revise their beliefs and hence in the limit their posteriors are the same as their priors. The limiting belief hierarchies will be common knowledge but will not be the same! (We will show in the next section that common priors are required for them to be the same.)

8.iv. Agreeing To Disagree in the Limit. Since at each n>1 agents observe the announcement $\delta(\tau_{n-1})$, this will be $\Im_{in}$-measureable for each $n > 1$ and $i\epsilon I$. Hence the sequence across time of these announcements, $\delta^\infty(\tau^\infty)\equiv\{\delta(\tau_n)\}^\infty_{n=1}$, is "known at date ∞ under $\Im_{i\infty}$"; or, more precisely, it is measureable with respect to the σ-field $\Im_{i\infty}$, for each $i\epsilon I$.

Now, as of "date ∞", i.e., under $\Im_{i\infty}$, each agent $i\epsilon I$ would have information from three sources: (a) Observation of the announcements $\delta^{\infty}(\tau^{\infty}) \equiv \{\delta(\tau_n)\}_{n=1}^{\infty}$; (b) knowledge its own type, $\tau_{i0} \epsilon T_i$; and (c) any supplementary private information from the observations $\{z_{in}\}_{n=1}^{\infty}$ not already accounted for in the announcements.

Suppose condition (GH) holds, and let the measure μ_i over the sample space Ω be agent i's prior subjective beliefs before observation of it's type as defined in eq. (7.5). At "date ∞" each agent $i \epsilon I$ will have beliefs represented by $\mu_i(. | \tau_{i0}, \mathfrak{F}_{i\infty})$. Consider an outside observer with the same prior μ_i but who only observes on the sample path of the announcements, $\delta^{\infty}(\tau^{\infty})$, and nothing else; in particular such an outside observer does not know i's type or any other private observations of agent i. At "date ∞" the outside observer has beliefs represented by $\mu_i(. | \delta^{\infty}(\tau^{\infty}))$.

Suppose that the announcements are of the form of any one of examples (a)-(c) in section 8.i above. We now show that on each sample path such an outside observer will make the same announcements as agent i. In particular we have the following:

Proposition 8.4 Fix any $i \epsilon I$. The following is true on a set of sample paths, $\omega \epsilon \Omega$, with μ_i probability one. If the announcement is of the type (a), (b) or (c) respectively, then,

(a) $\mu_i(E | \tau_{i0}, \mathfrak{F}_{i\infty}) = \mu_i(E | \delta^{\infty}(\tau^{\infty}))$; or,

(b) $\text{Marg}_{\theta}\ \mu_i(. | \tau_{i0}, \mathfrak{F}_{i\infty}) = \text{Marg}_{\theta}\ \mu_i(. | \delta^{\infty}(\tau^{\infty}))$; or,

(c) $\int \theta \mu_i(d\theta | \tau_{i0}, \mathfrak{F}_{i\infty}) = \int \theta \mu_i(d\theta | \delta^{\infty}(\tau^{\infty}))$, respectively.

The above proposition states that the limiting announcements of agents equals the announcements that would be made by an outside observer who has only information on those announcements and nothing else. An immediate corollary of the above result is the following: Suppose agents have a common prior, (i.e., suppose Harsanyi condition holds, and not it's generalization, condition (GH), so that $\mu_i = \mu_j$ for all i and j in I). Then each then each agent's limiting announcement must be the same. Of course from Proposition 8.2 we know that the limiting actions will be common knowledge. Hence we have:

Corollary 8.5. (Under Common Priors Agents Will Not Agree to Disagree!) Suppose that the Harsanyi Common Prior assumption holds, and in particular suppose that Condition (GH) holds with $\pi_i = \pi_j = \pi$ for all i and j in I. Suppose further that the announcement function one of examples (a)-(c) above. Then for μ_π almost every ω, agents limit announcements, $\{\delta_i(\tau_{i\infty})\}_{i\epsilon I}$, are common knowledge at (ω,∞) and are the same (i.e., $\delta_i(\tau_{i\infty})=\delta_j(\tau_{j\infty})$ for all $i,j\epsilon I$).

Suppose that the announcement function is of the form $\delta(\tau_n) =\Sigma_{i\epsilon I}f_i(\delta_i(\tau_{in}))$ where the set of functions $\{\delta_i\}_{i\epsilon I}$ is of the form of one of examples (a) or (c) and where the functions $\{f_i\}_{i\epsilon I}$ are each monotone real-valued functions. Then using the arguments of Nielson, et. al. (1990) it is easy to show that corollary 8.5 holds for this case.

8.v. On Partitions, Meets and Joins. Proposition 8.4 answers the question, "What happens to the Aumann (1976) Agreeing-To-Disagree" result when agents do **NOT** have common priors?" The result is however stated a very specific and dynamic setup. We now restate the classic Aumann (1976) to shed more light on the results of the previous sub-section.

In the Aumann framework there is a finite collection of agents, $i \epsilon I$, each with an information field, $\Im_i$, generated by a **finite** partition; (i.e., $\Im_i$ has finitely many elements). Each agent $i \epsilon I$ has a prior probability μ_i over the probability space $(\Omega, \Im_i)$. Recall that the meet, $\bigwedge_{i \epsilon I} \Im_i$, of the σ-fields is the finest σ-field that simultaneously coarsens each of them; the join, $V_{i \epsilon I} \Im_i$, is the coarsest common refinement. Assume μ_i assigns strictly positive probability to each element of the join.

In Aumann (1976) an event is common knowledge at $\omega \epsilon \Omega$ if and only if it contains the element of the meet at ω. With this it is easy to show that the original Aumann result may be restated as follows:

Proposition 8.6. (Aumann (1976)): (a) If posterior probability of some agent $i \epsilon I$, $\mu_i(. | \Im_i)$, is common knowledge at some $\omega \epsilon \Omega$, then the posterior must be the same as the prior probability conditional on the common element of the meet at ω; i.e., $\mu_i(. | \Im_i) = \mu_i(. | \bigwedge_{j \epsilon I} \Im_j)$.
(b) If each agent's posterior probability is common knowledge at $\omega \epsilon \Omega$, and, in addition, the agents have common priors (i.e., $\mu_i = \mu_j$ for all i and j in I), then those posteriors must be the same at

ω; i.e., $\mu_i(.|\mathfrak{F}_i) = \mu_j(.|\mathfrak{F}_j)$ for all $i,j \epsilon I$.

The conclusion of the original Aumann (1976) result was that under common priors, if the posteriors are common knowledge they must be the same. The above restatement shows what happens when the priors are not "common" (i.e., not the same): If the posteriors are common knowledge then the posteriors equal the prior conditional on the common element of the meet.

Set $\mathfrak{F}_i$ in Proposition 8.6 equal $\sigma(\{\tau_{i0},\mathfrak{F}_{i\infty}\})$, the σ-field generated by knowledge of agent i's date 0 type and limiting information. The sample path of the announcements of agents, $\delta^\infty(\tau^\infty)$, was shown in section 8.iii. to be common knowledge in the sense of definition 7.2. In a finite partitions world, this is equivalent to common knowledge in the sense of Aumann (1976).

Propostion 8.6(a) and (b) are then equivalent to Proposition 8.4(b) and Corollary 8.5 respectively. Indeed, via very simple modifications in the statement and proof of Propositon 8.6, we may obtain the analagous expressions for Proposition 8.4(a) and (c). The interpretation of Proposition 8.4. then is that if the priors are not common then the posteriors over θ (or more precisely the expressions in Proposition 8.4(a)-(c)) equal the prior conditional on the common observations, $\delta^\infty(\tau^\infty)$. Corollary 8.5 then says that Agreeing to Disagree will not occur if priors are "common."

9. Appendix. The Proofs.

Proof of Lemma 7.4: Let (ω,n) be given. Suppose that i knows X at (ω,n). Define $x=X(\omega)$. Fix any set A in U(S). Using definition 7.2, if $x \in A$, $\mu_{in}(\{\omega' \in \Omega \text{ s.t. } X(\omega') \in A\})(\omega) \geq \mu_{in}(\{\omega' \in \Omega \text{ s.t. } X(\omega')=x\})(\omega) = 1_{\{X(\omega)=x\}} = 1$ from which eq. (7.4) follows; and if $x \in S-A$ (the complement of A in S), $\mu_{in}(\{\omega' \in \Omega \text{ s.t. } X(\omega') \in A\})(\omega) \leq \mu_{in}(\{X(\omega') \in S-\{x\}\})(\omega) = 1 - 1_{\{X(\omega)=x\}} = 0$ from which eq. (7.4) follows.

Next suppose that eq. (7.4) holds at (ω,n). Define $x=X(\omega)$. Let S' denote a separant set of the separable space S. One can find a sequence of points $\{x_r\}^{\infty}_{r=1}$ in S' converging to x. One can also find a sequence of integers, n_r for each $r=1,2,\ldots,$ which converge to infinity and are such that for each r sufficiently large, x lies in the open ball $A_r \equiv B(x_r,1/n_r)$ with center x_r and radius $1/n_r$.

Define $A^M \equiv \bigcap^M_{r=1} A_r$. Then A^M monotonically decreases to $\{x\}$. Now by (7.4), for each $r=1,2,\ldots,$ since $X(\omega)=x \in A_r$ and $A_r \in U(S)$, $\mu_{in}(\{\omega' \in \Omega \text{ s.t } X(\omega') \in A_r\})(\omega) = 1$. This in turn implies that for each $M= 1,2,\ldots,$ $\mu_{in}(\{\omega' \in \Omega \text{ s.t } X(\omega') \in A^M\})(\omega) = 1$. Taking limits as $M \to \infty$ then implies that $\mu_{in}(\{\omega' \in \Omega \text{ s.t } X(\omega') \in \{x\}\})(\omega) = 1$ so (7.2) holds for $x=X(\omega)$. This in turn implies that the conditional probability $\mu_{in}(\cdot)(\omega)$ assigns probability one to X taking the value $x=X(\omega)$ at ω. Hence $\mu_{in}(\cdot)(\omega)$ assigns probability zero to any other value of x. In particular, (7.2) holds for all $x \in S$.//

Proof of Proposition 7.5: Under the hypothesis of the Proposition, $\mu_{\tau}(K^1{}_nD) = 1$. We proceed inductively. Suppose that for some r

=1,2,..., we have shown that $\mu_\pi(K^r{}_n D) = 1$. Then from the Mutual absolute continuity (M.A.C.) of μ_i and μ_π, $\mu_i(K^r{}_n D) = 1$. This in turn implies $\mu_i(K^r{}_n D | \tau_{i0}, \mathfrak{F}_{in})(\omega) = 1$, for μ_i-a.e.ω. However this latter equality is the definition of "i knows $K^r{}_n D$ at (ω,n)." Hence, $\mu_i(K^1{}_{in}(K^r{}_n D)) = 1$. The M.A.C. of μ_π and μ_i then implies $\mu_\pi(K^1{}_{in}(K^r{}_n D)) = 1$. Since this is true for all $i \epsilon I$, and I is finite we conclude that $\mu_\pi(K_n^{r+1} D) = 1$. Hence by induction we obtain that $\mu_\pi(K^r{}_n D) = 1$ for r. This in turn implies that $\mu_\pi(\bigcap^\infty{}_{r=1} K^r{}_n D) = 1$ from which the Proposition follows.//

Proof of Corollary 7.6: Fix any $i \epsilon I$. Under the hypothesis of this corollary, $\mu_\pi(K^1{}_n X) = 1$; so from the M.A.C. of μ_i and μ_π, $\mu_i(K^1{}_n X) = 1$. This implies, from the definition of a conditional probability, that $\mu_i(K^1{}_n X | \tau_{i0}, \mathfrak{F}_{in})(\omega) = 1$ for μ_i-a.e. ω. Hence $\mu_i(K^1{}_{in}(K^1{}_n X)) = 1$. From the M.A.C. property we have $\mu_\pi(K^1{}_{in}(K^1{}_n X)) = 1$. Since the set I is finite we conclude that $\mu_\pi(K^1{}_n(K^1{}_n X)) = 1$. In particular, for μ_π-a.e. ω, each agent knows the set $K^1{}_n X$. Proposition 7.5 in turn implies that $K^1{}_n X$ is common knowledge from which the corollary follows.//

Proof of Proposition 7.7: Fix any set A in U(S) (recall definition 7.3). Since for each $i \epsilon I$, X is $\mathfrak{F}_{in}$ measureable $\mu_i(\{\omega' \epsilon \Omega \text{ s.t. } X(\omega') \epsilon A\} | \tau_{i0}, \mathfrak{F}_{in})(\omega) = 1_{\{X(\omega) \epsilon A\}}$ for μ_i-a.e. $\omega \epsilon \Omega$. Since U(S) is countable, this holds for all sets $A \epsilon U(S)$ (simulataneously) for μ_i-a.e. ω. From Lemma 7.4 this is equivalent to "i knows X at (ω,n)." Hence $\mu_i(K^1{}_{in} X)=1$. From the M.A.C. condition this implies

$\mu_\pi(K^1{}_{in}X)=1$. Since I is finite we obtain $\mu_\pi(K^1{}_nX)=1$. An application corollary 7.6 then implies conclusion of this proposition. //

Proof of Corollary 7.8: Define X'_∞ to be equal to X_∞ on the set where $\lim_{n\to\infty}X_n$ exists and equal to zero where it does not exist. Since X_n is $\mathfrak{F}_{in}$ measureable for each $i\epsilon I$ it is easy to show that X'_∞ is $\mathfrak{F}_{i\infty}$-measureable. Proposition 7.7 then implies that for μ_π-a.e. ω, X'_∞ is common knowledge at (ω,∞). Since $X'_\infty = X_\infty$, μ_π-a.e., this implies the conclusion of the corollary. //

Proof of Proposition 8.1: When the announcement functions are continuous the result follows trivially from Theorem 6.1. The announcement function in example (b) is continuous, since it is in that case the projection map onto the first coordinate. The announcement function in part (c) is continuous because it is the integral on a compact, and hence bounded, set.

In example (a) the announcement function is not necessarily continuous. However, we may show directly in that case the required continuity. Indeed, fix any $i\epsilon I$ and $\tau_{i0}\epsilon T_i$. Observe that the sequence of posterior probabilities of E, $\{\mu_i(E|\tau_{i0},\mathfrak{F}_{in})\}^\infty_{n=1}$, is a bounded martingale sequence on the probability space $(\Omega,B(\Omega),\mu_i(.|\tau_{i0}))$. From the Martingale Convergence Theorem those posterior probabilities converge with $\mu_i(.|\tau_{i0})$ probability one. Since this is true for each τ_{i0} we obtain the convergence with μ_i probability one. From the mutual absolute continuity of μ_i and μ_π this is true with μ_π probability one. Since the set of agents is finite, this is true for each $i \in I$ (simultaneously) with μ_π probability one, from which the convergence of $\delta_i(\tau_{in})$ in example (b) follows. //

Proof of Proposition 8.4: (a) We are given the subset E of Θ. For ease of exposition we write $\mu_i(E)$ when we mean $\mu_i(\{\theta \in E\})$. Since $\delta^\infty(\tau^\infty)$ is measureable with respect to the information σ-field $\Im_{i\infty}$, $\mu_i(E|\tau_{i0}, \Im_{i\infty}) = \mu_i(E|\tau_{i0}, \Im_{i\infty}, \delta^\infty(\tau^\infty))$. From the convergence result of Proposition 8.1, $\mu_i(E|\tau_{i0}, \Im_{i\infty}, \delta^\infty(\tau^\infty)) \equiv \delta_i(\tau_{i\infty}) = \lim_{n\to\infty} \delta_i(\tau_{in})$. Hence

$$\mu_i(E|\tau_{i0}, \Im_{i\infty}) = \lim_{n\to\infty} \delta_i(\tau_{in}) \tag{8.6}$$

Now, $\mu_i(E|\delta^\infty(\tau^\infty)) = \int \mu_i(E|\tau_{i0}, \Im_{i\infty}, \delta^\infty(\tau^\infty)) d\mu_i(.|\delta^\infty(\tau^\infty))$. So integrating eq. (8.6) with respect to the probability $\mu_i(.|\delta^\infty(\tau^\infty))$ and noting that the limit in eq. (8.6) is measureable with respect to σ-field generated by $\delta^\infty(\tau^\infty)$, we obtain $\mu_i(E|\delta^\infty(\tau^\infty)) = \lim_{n\to\infty} \delta_i(\tau_{in})$. Using this in (8.6) and noting that all of the above statements hold μ_i-a.e. results in part (a).

(b) The arguments used in the proof of (a) shows that for each fixed subset D of Θ, $\mu_i(D|\tau_{i0}, \Im_{i\infty}) = \mu_i(D|\delta^\infty(\tau^\infty))$, μ_i-a.e. This equality can be made to hold for each set D in the class $U(\Theta)$ of definition 7.3 (simultaneously) on a set of sample paths with μ_i-probability one. It is easy to show that if two (Borel) probability measures on a metric space S agree on each element of U(S) in definition 7.3. then the two measures must be the same. Hence with μ_i-probability one, when restricted to Θ, $\mu_i(.|\tau_{i0}, \Im_{i\infty}) = \mu_i(.|\delta^\infty(\tau^\infty))$, and (b) follows.

(c) The proof of (c) is similar to (a) so is omitted. //

Proof of Proposition 8.6. Suppose at some $\omega \epsilon \Omega$, the posterior of agent i, $\mu'_i \equiv \mu_i(.|\mathfrak{F}_i)$, is common knowledge. Then from Aumann (1976) we know that the event that the posterior distribution takes the value μ'_i must contain the element of the meet at ω, $M \epsilon \bigwedge_{i \epsilon I} \mathfrak{F}_i$. Now, M is the finite union of elements in $\mathfrak{F}_i$, $M = \bigcup^R_{r=1} M_r$. On each such element i's posterior is μ'_i. Fix any subset D of Ω. Then by Bayes' rule, $\mu_i(D \cap M_r) = \mu'_i(D)\mu_i(M_r)$ for each r. By summation over r, $\mu_i(D \cap M) = \mu'_i(D)\mu_i(M)$; hence using Bayes' rule again, $\mu'_i(D) = \mu_i(D|M)$ from which part (a) follows. Part (b) is an immediate corollary of (a). //

Footnotes.

[1]In the framework of the model we study, our mutual absolute continuity assumption is however much weaker than that used by Blackwell and Dubins (1963) and Kalai and Lehrer (1990). See Nyarko (1991c) for details.

References.

Aumann, R. (1976):"Agreeing to Disagree," Annals of Statistics, 4, 1236-1239.

Billingsley, P. (1968): Convergence of Probability Measures, Wiley, New York.

Blackwell, D. and L. Dubins (1963):"Merging of Opinions with Increasing Information," Annals of Mathematical Statistics, 38, 882-886.

--------------------------- (1975):"On Existence and Non-existence of Proper, Regular Conditional Probabilities," the Annals of Probability, 3, 741-752.

Brandenburger, A. and E. Dekel (1987):"Common Knowledge with Probability 1," Journal of Mathematical Economics, 16, 237-245.

Geanakoplos, J. and H. Polemarcharkis (1982): "We Can't Disagree Forever," Journal of Economic Theory, 28, 192-200.

Harsanyi, J.C. (1967,1968): "Games with Incomplete Information Played by Bayesian Players," Parts I,II,III, Management Science, vol. 14, 3,5,7., pp. 159-182, 320-334, 486-502.

Laha, R.G. and V.K. Rohatgi (1979): Probability Theory, John Wiley and Sons, New York.

Kalai, E. and E. Lehrer (1990): "Rational Learning Leads to Nash Equilibrium," Discussion Paper No. 895, CMSEMS, Northwestern University.

McKelvey, R. and T. Page (1986):"Common Knowledge, Consensus, and Aggregate Information," Econometrica, 54, 109-127.

Mertens, J.-F., and S. Zamir (1985): " Formalization of Bayesian Analysis for Games with Incomplete Information," International Journal of Game Theory, 14:1-29.

Nielson, L. T. (1984):"Common Knowledge, Communication, and Convergence of beliefs," mathematical Social Sciences, 8, 1-14.

Nielson, L., A. Brandenburger, J. Geanakoplos, R. McKelvey and T. Page (1990):"Common Knowledge of an Aggregate of Expectations," Econometrica, 58, 1235-40.

Nyarko, Y. (1991a):"Learning in Mis-Specified Models and the Possibility of Cycles," Journal of Economic Theory, (forthcoming).

----------- (1991b):"Most Games Violate The Harsanyi Condition," C.V. Working Paper No. 91-39, New York University.

----------- (1991c):"The Convergence of Bayesian Belief Hierarchy," C.V. Working Paper, New York University.

Parthasarathy, K.R. (1967):"Probability Measures on Metric Spaces," Academic Press, New York.

Tan, T and S. da Costa Werlang (1988):"On Aumman's Notion of Common Knowledge -- An alternative Approach," CARESS Working Paper No. 88-09, University of Pennsylvannia.