Knowledge and Equilibrium

Jonathan Levin

April 2006

These notes develop a formal model of knowledge. We use this model to prove the "Agreement" and "No Trade" Theorems and to investigate the reasoning requirements implicit in different solution concepts. The notes follow Osborne and Rubinstein (1994, ch. 5). For an outstanding survey that covers much the same material, see Samuelson (2004).

1 A Model of Knowledge

The basic model we consider works as follows. There are a set of states Ω, one of which is true. For each state $\omega \in \Omega$, and a given agent, there is a set of states $h(\omega)$ that the agent considers possible when the actual state is ω. Finally, we will say that an agent *knows* E if E obtains at all the states that the agent believes are possible. These basic definitions will provide a language for talking about reasoning and the implications of reasoning.

States. The starting point is a set of states Ω. We can vary our interpretation of this set depending on the problem. In standard decision theory, a state describes contingencies that are relevant for a particular decision. In game theory, a state is sometimes viewed as a *complete* description of the world, including not only an agent's information and beliefs, but also his behavior.

Information. We describe an agent's knowledge in each state using an *information function.*

Definition 1 *An* ***information function*** *for Ω is a function h that associates with each state $\omega \in \Omega$ a nonempty subset $h(\omega)$ of Ω.*

Thus $h(\omega)$ is the set of states the agent believes to be possible at ω.

Definition 2 *An information function is* ***partitional*** *if there is some partition of Ω such that for any $\omega \in \Omega$, $h(\omega)$ is the element of the partition that contains ω.*

It is straightforward to see that an information function is partitional if and only if it satisfies the following two properties:

P1 $\omega \in h(\omega)$ for every $\omega \in \Omega$.

P2 If $\omega' \in h(\omega)$, then $h(\omega') = h(\omega)$.

Given some state ω, Property P1 says that the agent is not convinced that the state is not ω. Property P2 says that if ω' is also deemed possible, then the set of states that would be deemed possible were the state actually ω' must be the same as those currently deemed possible at ω.

Example 1 Suppose $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and that the agent's partition is $\{\{\omega_1, \omega_2\}, \{\omega_3\}\{\omega_4\}\}$. Then $h(\omega_3) = \{\omega_3\}$, while $h(\omega_1) = \{\omega_1, \omega_2\}$.

Example 2 Suppose $\Omega = \{\omega_1, \omega_2\}$, $h(\omega_1) = \{\omega_1\}$ but $h(\omega_2) = \{\omega_1, \omega_2\}$. Here h is not partitional.

Knowledge. We refer to a set of states $E \subset \Omega$ as an *event*. If $h(\omega) \subset E$, then in state ω, the agent views $\neg E$ as impossible. Hence we say that the agent *knows* E. We define the agent's *knowledge function* K by:

$$K(E) = \{\omega \in \Omega : h(\omega) \subset E\}.$$

Thus, $K(E)$ is the set of states at which the agent knows E.

Example 1, cont. In our first example, suppose $E = \{\omega_3\}$. Then $K(E) = \{\omega_3\}$. Similarly, $K(\{\omega_3, \omega_4\}) = \{\omega_3, \omega_4\}$ and $K(\{\omega_1, \omega_3\}) = \{\omega_3\}$.

Example 2, cont. In our second example, $K(\{\omega_1\}) = \{\omega_1\}$, $K(\{\omega_2\}) = \emptyset$ and $K(\{\omega_1, \omega_2\}) = \{\omega_1, \omega_2\}$.

Now, notice that for every state $\omega \in \Omega$, we have $h(\omega) \subset \Omega$. Therefore it follows that:

K1 (Axiom of Awareness) $K(\Omega) = \Omega$.

That is, regardless of the actual state, the agent knows that he is in some state. Equivalently, the agent can identify (is aware of) the set of possible states.

A second property of knowledge functions derived from information functions is that:

K2 $K(E) \cap K(F) = K(E \cap F)$.

Property K2 says that if the agent knows E and knows F, then he knows $E \cap F$. An implication of this property is that

$$E \subset F \Rightarrow K(E) \subset K(F),$$

or in other words, if F occurs whenever E occurs, then knowing F means knowing E as well. To see why this additional property holds, suppose $E \subset F$. Then $E = E \cap F$, so $K(E) = K(E \cap F)$. Applying K2 implies that $K(E) = K(E) \cap K(F)$, from which the result follows.

If the information function satisfies P1, the knowledge function also satisfies a third property:

K3 (Axiom of Knowledge) $K(E) \subset E$.

This says that if the agent knows E, then E must have occurred — the agent cannot know something that is false. This is a fairly strong property if you stop to think about it.

Finally, if the information function is partitional (i.e. satisfies both P1 and P2), the knowledge function satisfies two further properties:

K4 (Axiom of Transparency) $K(E) \subset K(K(E))$

Property K4 says that if the agent knows E, then he knows that he knows E. To see this, note that if h is partitional, then $K(E)$ is the union of all partition elements that are subsets of E. Moreover, if F is any union of partition elements $K(F) = F$ (so actually $K(E) = K(K(E))$).

K5 (Axiom of Wisdom) $\Omega \backslash K(E) \subset K(\Omega \backslash K(E))$.

Property K5 states the opposite: if the agent does not know E, then he knows that he does not know E. This brings home the strong assumption in the model that the agent understands and is aware of all possible states and can reason based on states that *might have* occurred, not just those that actually do occur (see Geanakoplos, 1994 for discussion).

We started with a definition of knowledge functions based on information functions, and derived K1-K5 as implications of the definition and the properties of partitional information. As it happens, it is also possible to go the other way because K1-K5 completely characterize knowledge functions. That is, Bacharach (1985) has shown that if we start with a set of states Ω and a function $K : \Omega \to \Omega$, then if K satisfies K1-K5 it is possible to place a partition on Ω to characterize the agent's information.

2 Common Knowledge

Suppose there are I agents with partitional information functions $h_1, ..., h_I$ and associated knowledge functions $K_1, ..., K_I$. We say that an event $E \subset \Omega$ is *mutual knowledge* in state ω if it is known to all agents, i.e. if $\omega \in K_1(E) \cap K_2(E) \cap ... \cap K_I(E) \equiv K^1(E)$. An event E is *common knowledge* in state ω if it is know to everyone, everyone knows this, and so on.

Definition 3 *An event $E \subset \Omega$ is* ***common knowledge in state*** *ω if $\omega \in K^1(E) \cap K^1K^1(E) \cap .....$*

A definition of common knowledge also can be stated in terms of information functions. Let us say that an event F is *self-evident* if for all $\omega \in F$ and $i = 1, ..., I$, we have $h_i(\omega) \subset F$.

Definition 4 *An event $E \subset \Omega$ is* ***common knowledge*** *in state $\omega \in \Omega$ if there is a self-evident event F for which $\omega \in F \subset E$.*

Example 3 Suppose $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and there are two individuals with information partitions:

$$\begin{aligned} \mathcal{H}_1 &= \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4\}\} \\ \mathcal{H}_2 &= \{\{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4\}\} \end{aligned}$$

- Is the event $E = \{\omega_1, \omega_2\}$ ever common knowledge? Not according to the second definition, since E does not contain any self-evident event. Moreover, not according to the first either since, $K_1(E) = \{\omega_1\}$ implies that $K_2K_1(E) = \emptyset$. Note, however, that $K^1(E) = \{\omega_1\}$ so E is mutual knowledge at ω_1.

- Is $F = \{\omega_1, \omega_2, \omega_3\}$ ever common knowledge? Apparently F is common knowledge at any $\omega \in F$ since F is self-evident. Moreover, since $K_1(F) = F$ and $K_2(F) = F$, it easy to check the first definition as well.

Lemma 1 *The following are equivalent: (i) $K_i(E) = E$ for all i, (ii) E is self-evident, (iii) E is a union of members of the partition induced by h_i for all i.*

Proof. To see *(i)* and *(ii)* are equivalent, note that F is self-evident iff $F \subset K_i(F)$ for all i. By K4, $K_i(F) \subset F$ always, so F is self-evident iff $K_i(F) = F$ for all i. To see *(ii)* implies *(iii)*, note that if E is self-evident, then $\omega \in E$ implies $h_i(\omega) \subset E$, so $E = \cup_{\omega \in E} h_i(\omega)$ for all i. Finally *(iii)* implies *(i)* immediately. *Q.E.D.*

Proposition 1 *The two definitions of common knowledge are equivalent.*

Proof. Assume E is common knowledge at ω according to the first definition. Then $E \supset K^1(E) \supset K^1K^1(E) \supset ...$ and ω is a member of each of these sets, which are thus non-empty. Since Ω is finite, there is some set $F = K^1 \cdots K^1(E)$ for which $K_i(F) = F$ for all i. So this set F, with $\omega \in F \subset E$ is self-evident and E is ck by the second definition.

Assume E is common knowledge at ω according to the second definition. There is a self-evident event F with $\omega \in F \subset E$. Then $K_i(F) = F$ for all i. So $K^1(F) = F$, and $K^1(F)$ is self-evident. Iterating this argument, $K^1 \cdots K^1(F) = F$ and each is self-evident. Now $F \subset E$, so by K2, $F \subset K^1 \cdots K^1(E)$. Since $\omega \in F$, E is ck by the first definition. *Q.E.D.*

3 The Agreement Theorem

Aumann (1976) posed the following question: could two individuals who share the same prior ever agree to disagree? That is, if i and j share a common prior over states, could a state arise at which it was commonly known that i assigned probability η_i to some event, j assigned probability η_j to that same event and $\eta_i \neq \eta_j$. Aumann concluded that this sort of disagreement is impossible.

Formally, let p be a probability measure on Ω — interpreted as the agents' prior belief. For any state ω and event E, let $p(E|h_i(\omega))$ denote i's posterior belief, so $p(E|h_i(\omega))$ is obtained by Bayes' rule. The event that "i assigns probability η_i to E" is $\{\omega \in \Omega : p(E|h_i(\omega) = \eta_i\}$.

Proposition 2 *Suppose two agents have the same prior belief over a finite set of states Ω. If each agent's infomation function is partitional and it is common knowledge in some state $\omega \in \Omega$ that agent 1 assigns probability η_1 to some event E and agent 2 assigns probability η_2 to E, then $\eta_1 = \eta_2$.*

Proof. If the assumptions are satisfied, then there is some self-evident event F with $\omega \in F$ such that:

$$F \subset \{\omega' \in \Omega : p(E|h_1(\omega') = \eta_1\} \cap \{\omega' \in \Omega : p(E|h_2(\omega') = \eta_2\}$$

Moreover, F is a union of members of i's information partition. Since Ω is finite, so is the number of sets in each union — let $F = \cup_k A_k = \cup_k B_k$. Now, for any nonempty disjoint sets C, D with $p(E|C) = \eta_i$ and $p(E|D) = \eta_i$, we have $p(E|C \cup D) = \eta_i$. Since for each k, $p(E|A_k) = \eta_1$, then $p(E|F) = \eta_1$ and similarly $p(E|F) = p(E|B_k) = \eta_2$. *Q.E.D.*

4 The No-Trade Theorem

The agreement theorem underlies an important set of results that place limits on the trades that can occur in differential information models under the common prior assumption (cf. Kreps, 1977; Milgrom and Stokey, 1982; Tirole, 1982; Osborne and Wolinsky, 1990). These "no-trade" theorems state, in various ways, that rational risk-averse traders cannot take opposite sides on a purely speculative bet. To see the basic idea, note that for some agent 1 to make an even money bet that a coin will come up Heads, he must believe that $\Pr(Heads) > 1/2$. For some agent 2 to take the other side of this bet, he must believe that $\Pr(Heads) < 1/2$. Aumann's theorem says the bet cannot happen since these opposing beliefs would then be common knowledge!

To formalize this idea, suppose there are two agents. Let Ω be a set of states and X a set of consequences (trading outcomes). A contingent contract is a function mapping Ω into X. Let A be the space of contracts. Each agent has a utility function $u_i : X \times \Omega \to \mathbb{R}$. Let $U_i(a) = u_i(a(\omega), \omega)$ denote i's utility from contract a — $U_i(a)$ is a random variable that depends on the realization of ω. Let $\mathbb{E}[U_i(a)|H_i]$ denote i's expectation of $U_i(a)$ conditional on his information H_i.

The following result is close cousin to the agreement theorem.

Proposition 3 *Let ϕ be a random variable on Ω. If i and j have a common prior on Ω, it cannot be common knowledge between them that i's expectation of ϕ is strictly greater than j's expectation of ϕ.*

Proof. This is on the problem set. *Q.E.D.*

Now, let us say that a contingent contract b is *ex ante efficient* if there is no contract a such that, for both i, $\mathbb{E}[U_i(a)] > \mathbb{E}[U_i(b)]$. We now state Milgrom and Stokey's (1982) no trade theorem.

Proposition 4 *If a contingent contract b is ex ante efficient, then it cannot be common knowledge between the agents that every agent prefers contract a to contract b.*

The proof will follow the same lines as the proof of the Agreement Theorem.

Proof. The claim is that there cannot be a state ω, which occurs with positive probability, at which the set

$$E = \{\omega : \mathbb{E}[U_i(a)|h_i(\omega)] > \mathbb{E}[U_i(b)|h_i(\omega)] \text{ for all } i\}$$

if common knowledge. Suppose to the contrary that there was such a state ω and hence a self-evident set F such that $\omega \in F \subset E$. By the definition of a self-evident set, for all $\omega' \in F$ and all i, $h_i(\omega') \in F$. So for all $\omega' \in F$ and all i:

$$\mathbb{E}[U_i(a) - U_i(b)|h_i(\omega')] > 0.$$

Now, using the fact that i's information is partitional, we know that $F = h(\omega^1) \cup h(\omega^2) \cup ... \cup h(\omega^n)$ for some set of states $\omega^1, ..., \omega^n \in F$ (in fact, we can choose these states so that $h(\omega^k) \cap h(\omega^l) = \varnothing$). It follows that for all i:

$$\mathbb{E}[U_i(a) - U_i(b)|F] > 0.$$

In other words, contract a strictly Pareto dominates contract b conditional on the event F. But now we have a contradiction to the assumption that b is ex ante efficient, because it is possible to construct a better contract c by defining c to be equal to a for all states $\omega \in F$ and equal to b for all states $\omega \notin F$. *Q.E.D.*

This theorem has remarkable implications because it says that under conditions that are quite standard in economic modelling — a common prior, Bayesian rationality and common knowledge being reached of any trades — there cannot be purely speculative transactions. Trades can only occur when the allocation of resources is not Pareto optimal and they must have an efficiency rationale.

This result seems nearly impossible to square with the enormous volume of trade in real-world financial markets. This suggests weakening one or more of the assumptions.

1. A route often followed in finance is to introduce "behavioral" or "liquidity" traders who are not rational but trade anyway. This can do

two things. First, to the extent that these traders lose money, they create gains from trade for rational traders. Second the presence of such traders may mean that two rational traders can trade without reaching common knowledge that they are trading – because one or both may think her trading partner is non-rational.

2. A second route is to relax the rationality assumption more broadly, for instance positing that agents use simple rules of thumb to process information or are systematically optimistic or pessimistic in interpreting information.

3. A third possibility is to weaken the common prior assumption so that agents "agree to disagree". In fact, Feinberg (2000) has shown that the common prior assumption is equivalent to the no-trade theorem in the sense that if it fails, agents can find some purely speculative bet that they would all want to take. Morris (1994) characterizes the set of trades that can occur with heteregenous priors.

5 Speculative Trade with Heterogeneous Priors

When agents have differing prior beliefs, the no-trade theorem fails. Here we'll cover an interesting case of trade with differing priors, taken from Harrison and Kreps (1978). To do this, we'll depart from the formal notation and just look at a simple example with numbers.

Consider a market where a single asset is traded by two classes of investors with infinite wealth, and identical discount factor $\delta = 3/4$. In every period $t = 1, 2, ...$, the asset pays a dividend d_t of either zero or one, and then can be traded. Both types of investors perceive the stream of dividends to be a stationary markov process with state space $\{0, 1\}$. But they disagree over the transition probabilities. Specifically, investors of type $i = 1, 2$ belive the transition probabilies are (here the ij entry reflects the probability of transiting from state i to state j):

$$Q_1 = \begin{matrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{matrix} \qquad\qquad Q_2 = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} .$$

The first type of investor believes the returns are independent over time; the second type believes the returns are perfectly positively correlated over time. Suppose we compute the value to each type of investor of holding the

asset for the entire future, as a function of today's dividend:

$$\begin{array}{cc} v^1(0) = 3/2 & v^1(1) = 3/2 \\ v^2(0) = 0 & v^2(1) = 3 \end{array}.$$

Thus in periods when the dividend is low, the first type of investor, who believes returns are iid, has a higher belief about the asset's fundamental value. But in periods when the dividend is high, the second investor, who believes returns are persistent, believes the fundamental value is higher. This suggests the asset will be traded back and forth depending on the current period dividend. But at what price?

Harrison and Kreps show that for every state s, and time t, prices must satisfy:

$$p_t(s_t) = \delta \bullet \max_k \sum_{s_{t+1}} \left[d_{t+1}(s_{t+1}) + p_t(s_{t+1})\right] Q_k(s_t, s_{t+1}).$$

That is, given a common expectation about the price process $p_t(\bullet)$, the price in state s must equal the maximum expected return across investors of buying the asset and holding it for a single period. (Why just a single period? Think about the optimality principle of dynamic programming.)

Given the stationarity of our example, equilibrium asset prices will also be stationary — that is, the price will be the same in every period the asset returns zero and in every period it returns one. Combining this observation with the Harrison and Kreps characterization of prices, we have:

$$\begin{aligned} p(0) &= \frac{3}{4}\left[\frac{1}{2}p(0) + \frac{1}{2}\left(1 + p(1)\right)\right] \\ p(1) &= \frac{3}{4}\left[1 + p(1)\right]. \end{aligned}$$

Solving these equations, we arrive at:

$$p(0) = 12/5 \qquad p(1) = 3.$$

In the high dividend state, type two investors buy the asset and are willing to hold it forever, so the price equals their fundamental valuation. In the low dividend, state, however, the price exceeds the fundamental valuation of both types of investors. Why? Because the price is set by type one investors who buy the asset for *speculative* purposes — that is, with the intention of holding the asset only until a dividend is paid at which point they intend to immediately unload the asset onto a type two investor.

This example seems quite special, but it can be generalized considerably (see Harrison and Kreps, 1978 and Scheinkman and Xiong, 2003). Nevertheless, this line of research has languished until recently despite many interesting questions one might ask. For instance, one might wonder how investor learning affects speculative trade. Morris (1996) shows that there can still be a speculative premium as learning occurs even if ultimately everyone agrees. Fudenberg and Levine (2005) argue that for natural learning processes, the end-result will be a situation where there is no trade (roughly they argue that a no-trade result obtains for a particular form of self-confirming equilibrium). Finally, you might notice that in the Harrison and Kreps model, investors agree on the state-price density (i.e. the price process) despite not agreeing on the transition probabilities over states. Mordecai Kurz has written a series of papers, using a quite different formalization, that attempt to relax this by allowing for heteregeneous beliefs directly about prices.

6 Epistemic Conditions for Equilibrium

Fix a game $G = (I, \{S_i\}, \{u_i\})$. Let Ω be a set of states. Each state is a complete description of each player's knowledge, action and belief. Formally, each state $\omega \in \Omega$ specifies for each i,

- $h_i(\omega) \subset \Omega$, i's knowledge in state ω.
- $s_i(\omega) \in S_i$, i's pure strategy in state ω.
- $\mu_i(\omega) \in \Delta(S_{-i})$, i's belief about the actions of others (note that i may believe other players actions are correlated).

We assume that among the players, it is common knowledge that the game being played is G. Thus, we assume each player knows his strategy set and the strategy sets and payoffs of the other players. Nevertheless, the model can be extended to games of incomplete information, where the game (e.g. payoff functions) may be uncertain. We will also maintain the assumption that each agent's information function is partitional.

Our first result is that if, in some state, each player is rational, knows the other players' strategies, and has a belief consistent with his knowledge, then the strategy profile chosen in that state is a Nash equilibrium of G.

Proposition 5 *Suppose that in state $\omega \in \Omega$, each player i:*

(i) knows the others' actions: $h_i(\omega) \subset \{\omega' \in \Omega : s_{-i}(\omega') = s_{-i}(\omega)\}$.

(ii) has a belief consistent with this knowledge: supp$(\mu_i(\omega)) \subset \{s_{-i}(\omega') \in S_{-i} : \omega' \in h_i(\omega)\}$

(iii) is rational: $s_i(\omega)$ *is a best response to* $\mu_i(\omega)$*,*

Then $s(\omega)$ *is a pure strategy Nash equilibrium of* G*.*

Proof. By *(iii)*, $s_i(\omega)$ is a best response for i to his belief, which by *(ii)* and *(i)* assigns probability one to the profile $s_{-i}(\omega)$. *Q.E.D.*

Clearly the assumption that each player knows the strategy choices of others is quite strong. What if we relax the assumption that players actually know the strategy choices of others, and replace it with an assumption that players know the *beliefs* of others and also that the others are rational? With two players, these conditions imply a Nash equilibrium in beliefs.

Proposition 6 *Suppose that* $I = 2$ *and that in state* $\omega \in \Omega$*, each player* i*:*

(i) knows the other's belief: $h_i(\omega) \subset \{\omega' \in \Omega : \mu_j(\omega') = \mu_j(\omega)\}$*,*

(ii) has a belief consistent with this knowledge,

(iii) knows that the other is rational: for any $\omega' \in h_i(\omega)$*, the action* $s_j(\omega')$ *is a best response of* j *to* $\mu_j(\omega')$ *for* $j \neq i$*.*

Then the profile $\sigma = (\mu_2(\omega), \mu_1(\omega))$ *is a Nash equilibrium of* G*.*

Proof. Let $s_i \in S_i$ be in the support of $\mu_j(\omega)$. By (ii) there is a state $\omega' \in h_j(\omega)$ such that $s_i(\omega') = s_i$. By (iii), s_i must be a best response to $\mu_i(\omega')$, which by (i) is equal to $\mu_i(\omega)$. *Q.E.D.*

Interestingly, neither of these results requires that beliefs be derived from a common prior on Ω. Indeed, beliefs need only be consistent with a player's knowledge. It is also not crucial that the game be common knowledge. For the first result, each player need only know his own preferences and strategy set. For the second, the game needs to be mutual knowledge.

Aumann and Brandenburger (1995) show that with three or more players, stronger assumptions are needed to justify mixed strategy Nash equilibrium. The main issue is to ensure that players i and j have the *same* beliefs about k. To ensure this, Aumann and Brandenburger assume a common prior, mutual knowledge of rationality and *common knowledge* of beliefs —

common knowledge of beliefs and a common prior imply identical beliefs by the Agreement Theorem.

Finally, we show that common knowledge of rationality implies rationalizability as we earlier suggested.

Proposition 7 *Suppose that in state $\omega \in \Omega$ it is common knowledge that each player's belief is consistent with his knowledge and that each player is rational. Then the profile $s(\omega)$ is rationalizable.*

Proof (for $I = 2$). Let $F \ni \omega$ be a self-evident event such that for every $\omega' \in F$ and each i, $s_i(\omega')$ is a best response to $\mu_i(\omega')$ and $\mu_i(\omega')$ is consistent with i's knowledge at ω'. For each i, let $B_i = \{s_i(\omega') \in S_i : \omega' \in F\}$. If $\omega' \in F$, then $s_i(\omega')$ is a best response to $\mu_i(\omega')$, whose support is a subset of $\{s_j(\omega'') \in S_j : \omega'' \in h_i(\omega')\}$. Since F is self-evident, $h_i(\omega') \subset F$, so $\{s_j(\omega'') \in S_j : \omega'' \in h_i(\omega')\} \subset B_j$. Thus $B_1 \times B_2$ is a best-reply set containing $s(\omega)$. *Q.E.D.*

It is also possible to provide an epistemic characterization of correlated equilibrium. In contrast to the other characterizations, which are *local* in nature (i.e. they refer to stategies, beliefs and knowledge at a particular state), correlated equilibrium is justified by assuming that players have a common prior and that *at every state* players are rational (this implies ck of rationality at every state). To state the result, due to Aumann (1987), let $\{H_i\}$ denote the partition induced by h_i for each i.

Proposition 8 *Suppose that for all $\omega \in \Omega$, (i) all players are rational, (ii) each player's belief is derived from a common prior p on Ω such that $p(h_i(\omega)) > 0$ for all i, ω, (iii) for all i, $s_i : \Omega \to S_i$ is measurable with respect to i's information, then $(\Omega, \{H_i\}, p, s)$ is a correlated equilibrium.*

Proof. Follows immediately from the definition of CE. *Q.E.D.*

Note that while we earlier motivated CE with an explicit communication story, here the information structure is just there. Gul (1998) argues that this leads to a conceptual problem — if the information structure is "just there," it's hard to know what to make of the common prior assumption.

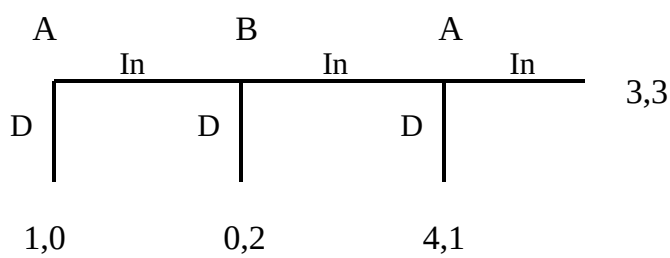
7 Comments

1. An idea we will return to throughout the term is that small departures from common knowledge can have a dramatic effect on the set of equilibria. In particular, even if each player is quite certain about the game

being player (i.e. the payoffs structure), even small uncertainty about other's information can eliminate equilibria that exist when payoffs are common knowledge. For one striking example of this, see Rubinstein (1989). Formally, the fact that small perturbations of the information structure can eliminate Nash equilibria occurs because the Nash equilibrium correspondence (mapping from the parameters of the game to the set of equilibrium strategies) is not lower semi-continuous.

2. Monderer and Samet (1989) define a version of "almost common knowledge" that has proved useful in some applications. There basic idea is that an event E is *common p-belief* if everyone assigns probability at least p to E, everyone assigns probability at least p to everyone assigning probability at least p to E and so on. Belief with probability 1 is then very close to knowledge (though not strictly speaking the same without some further assumptions), and common 1-belief the analogue of common knowledge.

3. Recent work has attempted to provide epistemic characterizations of dynamic equilibrium concepts such as subgame perfection, sequential equilibrium and so on. The difficulty is that rationality is a much more subtle issue in the extensive form. For example, consider the centipede game:



Centipede Game

Consider the following argument for why common knowledge of rationality should imply the backward induction solution that A play D immediately:

If A is rational, A will play D at the last node; if B knows A is rational, then B knows this; if B herself is rational, she must then play D at

the second to last node; if A knows B is rational, and that B knows that A is rational, then A knows that B will play D at the second to last node; thus, if A is rational, A must play D immediately.

The question, however, is *what will happen if* A plays In? At that point, how will B assess A's rationality? This puzzle has inspired much recent research (see Brandenburger, 2002, for a survey).

4. One interesting approach to these problems is Feinberg (2004), who proposes to treat each agent at each information set as essentially a separate individual (though all A's have the same preferences), with each individual having a frame of reference that assumes his node has been reached. Then B might still be confident of A's future rationality even if A's past rationality has been refuted. In Feinberg's account, common knowledge of rationality may actually contradict the logical structure of the game (the centipede game is such a game).

References

[1] Aumann, R. (1976) "Agreeing to Disagree," *Annals of Statistics.*

[2] Aumann, R. (1987) "Correlated Equilibrium as an Extension of Bayesian Rationality," *Econometrica*, 55, 1–18.

[3] Aumann, R. and A. Brandenburger (1995) "Epistemic Conditions for Nash Equilibrium," *Econometrica*, 63, 1161–1180.

[4] Bacharach, M. (1985) "Some Extensions to a Claim of Aumann in Axiomatic Model of Knowledge," *J. Econ. Theory*, 37, 167–190.

[5] Brandenburger, A. (2002) "The Power of Paradox," HBS Working Paper.

[6] Fagin, R., J. Halpern, Y. Moses, and M. Vardi (1995) *Reasoning about Knowledge*, Cambridge: MIT Press.

[7] Feinberg, Y. (2000) "Characterizing Common Priors in the Form of Posteriors," *J. Econ. Theory.*

[8] Feinberg, Y. (2004) "Subjective Reasoning – Dynamic Games," and "Subjective Reasoning – Solutions," *Games and Econ. Behavior.*

[9] Fudenberg, D. and D. Levine (2005) "Learning and Belief-Based Trade," Harvard Working Paper.

[10] Geanakoplos, J. (1994) "Game Theory without Partitions, with Applications to Speculation and Consensus" *J. Econ. Theory.*

[11] Gul, F. (1998) "A Comment on Aumann's Bayesian View," *Econometrica.*

[12] Harrison, M. and D. Kreps (1978) "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly J. Econ.*

[13] Kreps, D. (1977) "A Note on Fulfilled Expectations Equilibria," *J. Econ. Theory.*

[14] Milgrom, P. and N. Stokey (1982) "Information, Trade and Common Knowledge," *J. Econ. Theory.*

[15] Monderer, D. and D. Samet (1989) "Approximating Common Knowledge with Common Beliefs," *Games and Econ. Behavior,* 1, 170-190.

[16] Morris, S. (1994) "Trade with Asymmetric Information and Heterogeneous Priors," *Econometrica.*

[17] Morris, S. (1996) "Speculative Investor Behavior and Learning," *Quart. J. Econ.*, 111, 1111–1133.

[18] Rubinstein, A. (1989) "The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge'," *Amer. Econ. Rev.*

[19] Rubinstein, A. and A. Wolinsky (1990) "On the Logic of 'Agreeing to Disagree' Type Results," *J. Econ. Theory.*

[20] Samuelson, L. (2004) "Modeling Knowledge in Economic Analysis," *J. Econ. Lit.*, 42, 367–403.

[21] Scheinkman, J. and W. Xiong (2003) "Overconfidence and Speculative Bubbles," *J. Pol. Econ.*, 111, 1183–1219.

[22] Tirole, J. (1982) "On the Possibility of Speculation under Rational Expectations," *Econometrica.*