On Extensions to Aumann's Theory of Common Knowledge

Benjamin C. Wallace

April 2012

1 Introduction

In 1976, Robert Aumann published his widely influential mathematical formulation of the idea of "common knowledge" in [2]. This paper had a great impact in game theory, control theory, economics, and related fields because of the conclusions that it drew. Aumann's main theorem stated that if two agents began with identical prior beliefs (i.e. unbiased relative to each other) and their posterior beliefs following some event were common knowledge, then these posteriors would necessarily be identical. In Aumann's terms, two such agents could not "agree to disagree". Later, Geanakouplos and Polemarchakis [4] established a reasonable method by which such posteriors (or generally any kind of information) could "become" common knowledge. They showed, using Aumann's framework, that if two such agents communicated some piece of information back and forth repeatedly, then eventually this information would become common knowledge.

One reason Aumann's result is so intriguing is that the kind of behaviour he predicted is simply not consistent with real-world observations. Indeed, it is very often the case that two or more rationally-minded people who appear to both understand a common set of facts do agree to disagree. This discrepancy led to much analysis of Aumann's assumptions and to refinements of his results. For instance, one might suppose that the convergence of beliefs predicted in [4] might simply require too much computational power or too much information transmission. Indeed, this question was brought up by Aaronson [1]; however, he concluded that this was *not* the case.

Perhaps one of the more obvious differences between Aumann's model and the real world is the fact that the model was developed over countable probability spaces. In the present paper, we shall discuss an extension of this model to general state spaces put forward by Nielsen [7]. Nielsen gives a similar conclusion as Aaronson by showing that Aumann's theorem is not truly restricted by the assumption of countability. Nielsen also generalizes Geanakouplos and Polemarchakis's results.

In further support of the Aumann-Nielsen model, we will discuss a paper of Brandenburger and Dekel [3], in which they give a rather different—but perhaps intuitively more approachable—

definition of common knowledge. Their main result is that, under some relatively mild assumptions, their definition agrees with Nielsen's.

2 Common Knowledge in General State Spaces

2.1 General Setup

In all that follows, we take $(\Omega, \mathcal{F}, P)$ to be an ambient probability space; in talking about a "prior" we will mean the probability distribution P. Following [7], however, we will work in the "generalized probability space" over $(\Omega, \mathcal{F}, P)$. More precisely, we define the equivalence relation

$$A \simeq B \text{ iff } P(A \triangle B) = P((A \setminus B) \cup (B \setminus A)) = 0$$

over events $A, B \in \mathcal{F}$ and consider "generalized events" $[A] = \{B : B \simeq A\} \in \mathcal{F}/\simeq$, i.e. equivalence classes of events under $\simeq$. It is easily seen that under the obvious extensions of the subset relation $\subset$, union $\cup$, intersection $\cap$, and complement $\cdot^c$, the quotient $\mathcal{F}/\simeq$ satisfies the axioms of a σ-algebra. However, $\mathcal{F}/\simeq$ is *not*, strictly speaking, a σ-algebra, as it is not obviously related to any underlying state space[1]. Note also that P extends naturally to $\mathcal{F}/\simeq$ and continues to satisfy the properties of a probability measure. In order to make our discussion more transparent, we will denote generalized events $[A]$ by any of their representative $B \simeq A$, write $A \subseteq B$ to denote subset inclusion in $\mathcal{F}/\simeq$ (i.e. up to null differences), write $A \in F$ to mean $[A] \in \mathcal{F}/\simeq$, etc.

2.2 Information and Knowledge

In order to motivate Nielsen's definition of knowledge, it is useful to recall the information-centric interpretation of σ-algebras. We denote the "true state of the world" by $\omega \in \Omega$ and say that an event $A \in \mathcal{F}$ *occurs* when $\omega \in A$. The natural interpretation of subset inclusion $A \subseteq B$ is as the implication "if A occurs, then B must occur". We identify each agent $i \in I$—where I is a finite index set—with their *information σ-algebra* $\mathcal{F}^i \subseteq \mathcal{F}$, the collection of events that i has complete information about; that is, if an event $A \in \mathcal{F}^i$ occurs, then agent i knows that it has occurred. As a result, agent i will know all events containing A.

Definition (Nielsen)**.** If $\mathcal{G}$ is a family of sets, then we denote by $\bigcup \mathcal{G}$ the union of all sets in $\mathcal{G}$. We follow a similar convention for intersections. Define the event that i *knows* A to be the (generalized) event

$$K(\mathcal{F}^i, A) = \bigcup \{F \in \mathcal{F}^i : F \subseteq A\}.$$

[1] For instance, $\mathcal{F}/\simeq$ is not a σ-field over the power set 2^Ω, as $\mathcal{F}/\simeq$ does not have 2^Ω as an element. Halmos [5] refers to objects such as $F/\simeq$ as *Boolean* σ-algebras; these capture the algebraic structure of a σ-algebra without imposing restrictions on its content.

That is, $K(\mathcal{F}^i, A)$ occurs precisely when an event in $\mathcal{F}^i$ occurs that will lead i to "infer" that A has occurred.

When Ω is countable we can eliminate events of probability 0, thereby reducing $\mathcal{F}^i/\simeq$ to $\mathcal{F}^i$. In this case, the definition above essentially agrees with that in Aumann (which uses information *partitions* to represent information) once we notice that any σ-algebra over a countable space can be generated by a countable partition of the space and that this partition can be recovered from the σ-algebra.

We also have the following simple but important properties:

1. $K(\mathcal{F}^i, A) \in \mathcal{F}^i$;

2. $K(\mathcal{F}^i, A) \subseteq A$;

3. if $A \in \mathcal{F}^i$, then $K(\mathcal{F}^i, A) = A$.

The first of these states that the event "i knows A" is available information to i. In particular, this means that $K(\mathcal{F}^i, A)$ really is a generalized event (we hadn't actually shown this yet and this isn't obvious as the union involved may very well be uncountable). This follows from the fact that $\mathcal{F}/\simeq$ forms a separable metric space[2] under the metric $d(A, B) = P(A \triangle B)$.

The last two properties simply state that if A is known to occur, then it must occur and that the converse of this holds for those A which are available to agent i; these follow straight from the definition.

2.3 An Alternative Formulation

We now turn to an observation made in Brandenburger and Dekel [3]. In this section alone, we assume that for all $F \in \mathcal{F}^i$, $P(F|\mathcal{F}^i)(\omega) = 1_F(\omega)$ for each $\omega \in \Omega$ (this is called *properness* of the conditional probability $P(\cdot|\mathcal{F}^i)$). In other words, we ask that $\mathcal{F}^i$ know for certain whether its events have occurred or not. In particular, if $\omega \in K(\mathcal{F}^i, A)$, then $P(A|\mathcal{F}^i)(\omega) = 1$, i.e.

$$K(\mathcal{F}^i, A) \subseteq \{\omega \in \Omega : P(A|\mathcal{F}^i)(\omega) = 1\}. \tag{1}$$

Brandenburger and Dekel provide the following instructive example in which the converse of this does not hold.

Example 1 (Brandenburger/Dekel)**.** Let

$$\Omega = \{\omega_1, \omega_2, \omega_3\}, \quad \mathcal{F} = 2^\Omega, \quad \mathcal{F}^1 = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3\}, \Omega\},$$
$$P(\{\omega_1\}) = P(\{\omega_3\}) = \frac{1}{2}, \quad P(\{\omega_2\}) = 0, \quad A = \{\omega_1\} \in \mathcal{F}.$$

[2] Thanks to Prof. Yüksel for suggesting this possibility, which was confirmed in [5].

Then

$$P(A|\mathcal{F}^1)(\omega_1) = P(A|\mathcal{F}^1)(\omega_2) = 1,$$
$$P(A|\mathcal{F}^1)(\omega_3) = 0$$

but $\omega_1, \omega_2 \notin K(\mathcal{F}^1, A) = \emptyset$.

If we introduce $\{\omega_2\}$ into $\mathcal{F}^1$, then the converse of (1) *will* hold. Such a "completion" of $\mathcal{F}^1$ may be interpreted as allowing agent 1 to know that $\{\omega_2\}$ "won't happen"; hence, when $\{\omega_1, \omega_2\}$ occurs, agent 1 "infers" that it must be $\{\omega_1\}$ that has occurred. More precisely, if $\{\omega_2\} \in \mathcal{F}^1$, then $\{\omega_1\} \in \mathcal{F}^1$ and $P(A|\mathcal{F}^1)(\omega_2) = 0$ and $\omega_1 \in K(\mathcal{F}^1, A) = \{\omega_1\}$.

More generally, Aumann found that a sufficient condition for the converse of (1) to hold is that $\mathcal{F}^i$ be *posterior complete*, in the sense that it contain all $B \in \mathcal{F}$ with $P(B|\mathcal{F}^i)(\omega) = 0$ for all $\omega \in \Omega$.

Proposition 2.1 (Brandenburger/Dekel[3])**.** *If $P(\cdot|\mathcal{F}^i)$ is proper and $\mathcal{F}^i$ is posterior complete, then*

$$K(\mathcal{F}^i, A) = \{\omega \in \Omega : P(A|\mathcal{F}^i)(\omega) = 1\}.$$

So under some relatively mild assumptions, knowledge of an event by an agent can be characterized entirely in terms of the conditional probability assigned to that event by the agent. This characterization states that i knows A if A assigns probability 1 to A given the information $\mathcal{F}^i$.

2.4 Common Knowledge

All the following concepts will be defined in terms of the event $K(\mathcal{F}^i, A)$. Hence, the result of Brandenburger and Dekel carries over to them.

Definition (Nielsen)**.** We define the event that *i knows a random variable X* as

$$K(\mathcal{F}^i, X) = \bigcap \{K(\mathcal{F}^i, A) \cup K(\mathcal{F}^i, A^c) : A \in \sigma(X)\}.$$

So i knows X when i knows, for each $A \in \sigma(X)$, whether or not A occurs. Agent i need not know about events outside of $\sigma(X)$ as knowledge of a random variable should only involve knowledge of the value it takes on.

So far, we have not said what is meant by "common knowledge" in any of the papers we are discussing, all of which follow Aumann's intuitive definition of common knowledge. Aumann is *not* referring to, for instance, the event that a collection of agents all have knowledge of an

[3] In [3], the statement of this and the following result are made in terms of *common knowledge*. However, the proofs given are strong enough to apply the more primitive notion of *knowledge* of an event as defined above and it is in these terms that we state these results.

event, i.e. he does not mean the event $\bigcap_{i\in I} K(\mathcal{F}^i, A)$ for some A. Rather, what is being discussed is the event that all agents know A, all know that each other know A, all know that each other know that each other know A, etc.

Definition (Aumann, Nielsen)**.** We define the event that A is common knowledge as the generalized event

$$K(A) = \bigcap_{\substack{i_1,\ldots,i_n\in I\\ n\in\mathbb{N}}} K(\mathcal{F}^{i_1}, K(\mathcal{F}^{i_2}, \ldots, K(\mathcal{F}^{i_n}, A)\ldots)).$$

The event $K(X)$ that X is common knowledge has the same definition with A replaced by X.

Since the set of all finite sequences over a finite set is countable, this is indeed a generalized event. Nielsen goes on to generalize the following characterization of common knowledge, first proved by Aumann.

Proposition 2.2 (Aumann, Nielsen)**.** *The following two statements hold:*

1. $K(A) = K\left(\bigcap_{i\in I}\mathcal{F}^i, A\right)$;
2. $K(X) = K\left(\bigcap_{i\in I}\mathcal{F}^i, X\right)$.

In lattice-theoretic terms, the events that are common knowledge are those that contain an element of the *meet* (finest common coarsening) $\bigcap_i \mathcal{F}^i$ of the $\mathcal{F}^i$.

2.5 Main Results

As we mentioned in the introduction, Aumann's theorem on the impossibility of "agreeing to disagree" and Geanakouplos and Polemarchakis's result on the convergence of beliefs hold in Nielsen's generalized framework.

Theorem 2.3 (Aumann, Nielsen)**.** *For all $j, k \in I$,*

$$P\left(E[X|\mathcal{F}^j] = E[X|\mathcal{F}^k] \,\middle|\, \bigcap_i K(E[X|\mathcal{F}^i])\right) = 1.$$

That is, if the expectations $E[X|\mathcal{F}^i]$ are common knowledge, then they are all equal.

Theorem 2.4 (Geanakouplos/Polemarchakis, Nielsen)**.** *Let X be an integrable random variable, $\mathcal{F}^i_n$ a filtration for each $i \in I$, and suppose that for each n, there exists $m > n$ such that $P(K(\mathcal{F}^i_m, E[X|\mathcal{F}^j_n])) = 1$ for all $i, j \in I$. Then for all $i \in I$,*

$$E[X|\mathcal{F}^i_n] \xrightarrow{n\to\infty} E\left[X \,\middle|\, \bigcap_i \mathcal{F}^i_\infty\right], \qquad \textit{(a.s.)}$$

where $\mathcal{F}^i_\infty = \bigcup_n \mathcal{F}^i_n$.

The main assumption here is that the agents all eventually know each other's updated expectations $E[X|\mathcal{F}^j_n]$. Integrability of X is required in order to make use of a martingale convergence theorem. When these assumptions hold, the theorem states that all agents expectations of X converge to the expectation of X given everyone's eventual common information $\bigcap_i \mathcal{F}^i_\infty$. Unlike Geanakouplos and Polemarchakis's result however, this convergence is not guaranteed to occur in a finite number of steps. This is, of course, a consequence of abandoning countability.

3 Discussion

The generalizations provided by Nielsen motivate analogous generalizations of work based on Aumann's original paper. For instance, Aaronson's result on the computational and communication complexity involved in the convergence of beliefs is restricted to the discrete setting. Other aspects of this convergence are also interesting, particularly in the generalized setting presented above. As mentioned, loss of a countable state space results in loss of finite convergence time. The natural problem that presents itself then is that of determining rates of convergence.

Looking further, it could be of interest to undertake an analysis of different information-sharing patterns. Nielsen only assumes that the agents involved share their information eventually; this is understandable, as only the qualitative aspect of convergence is established. However, a quantitative analysis of convergence may very well depend on the way in which information is allowed to be transmitted. In this way, tighter links could be established with decentralized control, in which information patterns play an important role.

In [4], it is shown that there is a slight discrepancy between the "belief equilibrium" achieved through the communication of beliefs (processed information) and that achieved through the communication of raw data. That is, the equilibriums reached in each of these manners need not be equal. We term this discrepancy "slight" because it is also shown that they are equal with probability 1. A similar analysis could be extended to Nielsen's framework.

Perhaps one of the most significant questions that arose immediately after Aumann's publication was whether or not his model was "correct". Support for the model was provided by Milgrom [6], who derived Aumann's definition of common knowledge from an axiomatic foundation; it would be interesting to know whether such a foundation would extend to Nielsen's setting. As far as justifying and better understanding any mathematical model goes, it is always advantageous to have several alternate characterizations. In this regard, Brandenburger and Dekel are extremely successful at illuminating the idea of common knowledge.

Any mathematical model must ultimately undertake much scrutiny before it becomes well-accepted. Part of such scrutiny involves extending, refining, and specializing the model in various ways and observing the behaviors resulting from these changes. Nielsen's extension of Aumann's model, for instance, behaves as one would hope.

However, the limitations we mention above do not necessarily indicate weaknesses or flaws in Nielsen's paper, just as the restriction to countable spaces did not truly limit Aumann's work (as made evident by Nielsen). Rather, these simply indicate the diverse avenues of research opened up in [2]. The rigorous formulation of common knowledge in the more refined framework of generalized σ-algebras makes evident the power of—and may even help to illuminate—the measure-theoretic interpretation of information.

References

[1] Scott Aaronson. The complexity of agreement. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, STOC '05, pages 634–643, New York, NY, USA, 2005. ACM.

[2] Robert J. Aumann. Agreeing to disagree. *The Annals of Statistics*, 4(6):1236–1239, November 1976.

[3] Adam Brandenburger and Eddie Dekel. Common knowledge with probability 1. *Journal of Mathematical Economics*, 16(3):237 – 245, 1987.

[4] John D Geanakoplos and Heraklis M Polemarchakis. We can't disagree forever. *Journal of Economic Theory*, 28(1):192 – 200, 1982.

[5] Paul Richard Halmos. *Measure Theory*. Springer-Verlag New York Inc., 1974.

[6] Paul Milgrom. An axiomatic characterization of common knowledge. *Econometrica*, 49(1):219–222, January 1981.

[7] Lars Tyge Nielsen. Common knowledge, communication, and convergence of beliefs. *Mathematical Social Sciences*, 8(1):1 – 14, 1984.