Aumann's "No Agreement" Theorem Generalized

Matthias Hild
Christ's College, Cambridge

Richard Jeffrey
Department of Philosophy, Princeton University

Mathias Risse
Department of Philosophy, Princeton University

August 7, 1998

(For *The Logic of Strategy*, ed. C. Bicchieri, R. Jeffrey, B. Skyrms: OUP)

Abstract

The scope of Aumann's (1976) Theorem is needlessly limited by its restriction to Conditioning as the update rule. Here we prove the theorem in a more comprehensive framework, in which the evolution of probabilities is represented directly, without deriving new probabilities from new certainties. The framework allows arbitrary update rules subject only to Goldstein's requirement that current expectations agree with current expectations of future expectations.

Introduction

In "Agreeing to Disagree", Robert Aumann (1976) proves that a group of agents who once agreed about the probability of some proposition for which their current probabilities are common knowledge must still agree, even if those probabilities reflect disparate observations. Perhaps one saw that a card was red and another saw that it was a heart, so that, as far as that goes, their common prior probability of 1/52 for its being the Queen of hearts would change in the one case to 1/26, and in the other to 1/13. But if those are indeed their current probabilities, it cannot be the case that both know them, and both know that both know them, and so on.

In Aumann's framework, new probabilistic states of mind can only arise by conditioning old ones on new knowledge. In such a framework, current probabilities must derive from what is in effect knowledge, that is, true full belief. But here we derive Aumann's result from common knowledge of a probability, however arrived at. We work with possible worlds in which the agents' probabilities and their evolution are matters of fact, represented within the model (i.e., as proposed by Samet (1990, 205)).

Independence of particular update rules is a central feature of the new framework. But of course we need some constraint on how agents update their probabilities. For this we use Goldstein's requirement (1983) that current expectations of future expectations equal current expectations. This is the workhorse for our proof of the Generalized "No Agreement" Theorem.

Related projects proceed differently. Thus Samet generalizes the theorem for agents who condition on certainties that need not be true, provided only that all agents are certain of their certainties; and both Cave (1983) and Bacharach (1985) generalize it for agents whose common knowledge is not of probabilities but of decisions. But in each case the gain in generality remains within the confines of the conditioning framework, where new probabilities can only stem from new certainties.

1 Aumann's Theorem

Aumann's framework provides for a finite number of *agents*; call them $i = 1, \ldots, N$. These individuals are about to learn the answers to various multiple-choice questions—perhaps by making observations. The possible answers to agent i's question form a set $\mathcal{Q}_i$ of mutually exclusive, collectively exhaustive propositions. We think of propositions as subsets of a non-empty set Ω of *worlds*, representing all possibilities of interest for the problem at hand. Then $\mathcal{Q}_i$ is a *partition* of Ω: each world ω belongs to exactly one element of each $\mathcal{Q}_i$. We call that element '$\mathcal{Q}_i\omega$'.

Perhaps it is only certain propositions, certain subsets of Ω, that are of interest to the agents. Among them will be all propositions in any of the

N partitions, and perhaps other propositions as well. We will suppose that they form a σ-field $\mathcal{A}$.

The propositions agents know after learning the answers to their questions are the members of $\mathcal{A}$ that those answers imply: In world ω agent i knows A if and only if $\mathcal{Q}_i\omega \subseteq A \in \mathcal{A}$. Then for each i, Def. K, below, defines a knowledge operator K_i which, applied to any set $A \in \mathcal{A}$, yields the set $K_iA \in \mathcal{A}$ of worlds in which i knows A.

(Def. K) $$K_iA = \{\omega : \mathcal{Q}_i\omega \subseteq A\}$$

For each natural number n, Def. MK then defines an operator M_n, "n'th degree mutual knowledge": A is true, and everybody knows that, and everybody knows *that*, and so on—with n 'knows'. Finally, Def. CK defines *common knowledge*, κ, as mutual knowledge of all finite degrees:

(Def. MK) $$M_0A = A$$ $$M_{n+1}A = \bigcap_{i=1}^{N} K_iM_nA$$

(Def. CK) $$\kappa A = \bigcap_{n=0}^{\infty} M_nA$$

The key to the proof of Aumann's theorem is the following lemma.

Lemma 1.1 *Where something is common knowledge, everyone knows it is: If $\kappa A \neq \emptyset$ then κA is the union of some subset $\mathcal{D}_i$ of $\mathcal{Q}_i$.*

Proof. If $\omega \in \kappa A$ then by Def. CK and Def. M, $\omega \in K_iM_nA$ for all agents i and degrees n of mutual knowledge. Therefore by Def. K, $\mathcal{Q}_i\omega \subseteq M_nA$ for all n, and thus by Def. CK, $\mathcal{Q}_i\omega \subseteq \kappa A$. ■

The first hypothesis of Aumann's theorem 1.2, below, says that $\langle\Omega, \mathcal{A}, P\rangle$ is a probability space, that $\mathcal{A}$ includes each of the partitions $\mathcal{Q}_1, \ldots, \mathcal{Q}_N$ of Ω, and that each of those partitions is countable. In jargon: $\langle\langle\Omega, \mathcal{A}, P\rangle, \mathcal{Q}_1, \ldots, \mathcal{Q}_N\rangle$ is a *countable partition space.* $\mathcal{A}$ will then be closed under all of the operators K_i and M_n, and under κ.

P is the old probability measure that is common to all the agents. In Aumann's theorem we consider a hypothesis $H \in \mathcal{A}$ for which the various agents' probabilities are $q_1, \ldots, q_N$ after they condition P on the answers to their questions. The proposition $C \in \mathcal{A}$ identifies these probabilities:

(Def. C) $$C = \bigcap_{i=1}^{N} \{\omega : P(H|\mathcal{Q}_i\omega) = q_i\}$$

The second hypothesis says that the possibility of C's becoming common knowledge is not ruled out in advance: $P(\kappa C) \neq 0$.

Theorem 1.2 (Aumann's "No Agreement" Theorem)
If $\langle\langle\Omega, \mathcal{A}, P\rangle, \mathcal{Q}_1, \ldots, \mathcal{Q}_N\rangle$ is a countable partition space, and $P(\kappa C) > 0$, then $P(H|\kappa C) = q_1 = \cdots = q_N$.

Proof. $P(H|\kappa C) = \frac{P(H \cap \bigcup_{D\in\mathcal{D}_i} D)}{\sum_{D\in\mathcal{D}_i} P(D)} = \frac{\sum_{D\in\mathcal{D}_i} P(H|D)P(D)}{\sum_{D\in\mathcal{D}_i} P(D)} = \frac{\sum_{D\in\mathcal{D}_i} q_i P(D)}{\sum_{D\in\mathcal{D}_i} P(D)} = q_i$

The first equation is justified by Lemma 1.1; the second by the probability calculus; the third by Def. C; the fourth by independence of q_i from D. ■

This remarkable result is limited by the assumption that agents update only by conditioning on authoritative answers to questions. Before removing that limitation, we pause to see why conditioning is not the only way to update.

2 Conditioning Generalized

In Aumann's framework, agents update their probabilities by conditioning on countable partitions of Ω. This will soon be replaced by a framework which is independent of how agents update. But first we point out an intermediate generalization of Aumann's updating scheme, in which (1) partitions are replaced by more general structures, and (2) conditioning is replaced by a more general operation.

2.1 Sufficient Subfields

The partitions $\mathcal{Q}_i$ are meant to represent the possible answers to questions that agents might obtain, for example, by making observations. But even in in the simplest cases, where the most specific answers form a finite partition of Ω, observation may only identify the disjunction of two or more partition elements as true without identifying any one element as true. And in case the observation is made on the value of a continuous magnitude, where total precision is unattainable, the possible results will form no partition but a family of overlapping intervals.

In a generalization of Aumann's model covering such cases, partitions $\mathcal{Q}$ of Ω are replaced by sub-σ-fields $\mathcal{F}_i$ of $\mathcal{A}$, and the partition element $\mathcal{Q}_i\omega$ which answers agent i's question in world ω is replaced by the subset $\mathcal{F}_i\omega = \{A : \omega \in A \in \mathcal{F}_i\}$ of $\mathcal{A}$ consisting of the elements of $\mathcal{F}_i$ that are true in world ω. (Billingsley, 1995, 57.) In the special case where the subfield is atomic, there is an equivalent partition model in which the elements of the partition are the atoms of the subfield. Here the notion of sufficiency has a simple definition: *Sufficiency* of $\mathcal{Q}_i$ for a family of measures means that conditioning on any member of $\mathcal{Q}_i$ wipes out all differences between members of the family. Where the family is the pair $\{P, Q_i\}$, the definition comes to this:[1]

(Sufficiency) $Q_i(\cdot|A) = P(\cdot|A)$ for any $A \in \mathcal{Q}_i$

Problem: In world ω, how is agent i to update P to a suitable new probability measure Q_i, upon discovering which member of $\mathcal{Q}_i$ is true? The discovery is a matter of full belief in the true member of $\mathcal{Q}_i$:

(Certainty) $Q_i(\mathcal{Q}_i\omega) = 1$

This problem has no general solution.[2] But as is easily verified, it has the following solution if, and only if, Sufficiency is satisfied:

(Conditioning) $Q_i(\cdot) = P(\cdot|\mathcal{Q}_i\omega)$

2.2 Uncertain Information

Whether or not partitions are replaced by subfields, we can drop the certainty condition to obtain a way of updating on uncertain information that is more broadly applicable than conditioning. Thus, in the case of a countable partition model, observation under less than ideal conditions may change i's probabilities for elements of $\mathcal{Q}_i$ from those given by the common prior measure P to new values, given by some measure Q_i for which $Q_i(\mathcal{Q}_i\omega) < 1$ for all ω. If Sufficiency is satisfied, this is updating by *Generalized Conditioning*:[3]

(Generalized Conditioning) $Q_i(\cdot) = \sum_{A \in \mathcal{Q}_i} P(\cdot|A)Q_i(A)$

Generalized conditioning is not a universally applicable method of updating. Rather, it is applicable if and only if the Sufficiency condition is met. Where Sufficiency fails, other special conditions may hold, in the presence of which other methods of updating may be applicable. It would be a hopeless task to try to form an inventory of all updating methods and their conditions of applicability. Instead, we now turn to a revised framework, in which we can track flows of probability without reference to whatever update rules they follow.

[1] In general, $\mathcal{F}_i$ is a sufficient subfield iff $Q_i(\cdot \parallel \mathcal{F}_i) = P(\cdot \parallel \mathcal{F}_i)$. See Billingsley (1995, 450).

[2] In the playing card example at the beginning, certainty about the true member of the partition {*heart, diamond, club, spade*} brought with it certainty about the true member of the partition {*red, black*}, but conditioning on these certainties led to different probabilities (1/26, 1/13) for the card's being the Queen of hearts.

[3] This follows from Sufficiency by the Law of Total Probability. For fully generalized conditioning, i.e., on a sub–σ–field $\mathcal{S}_i$, the definition would be $Q_i(\cdot) = \int_\Omega P(\cdot \| \mathcal{S}_i) dQ_i = E(P(\cdot \| \mathcal{S}_i))$, where E is the updated expectation operator. (Generalized Conditioning is sometimes called 'Probability Kinematics'.) For references and further information see Diaconis and Zabell (1982) or Jeffrey (1992).

3 The Theorem Generalized

As before, $1, \ldots, N$ are the agents, Ω is a non–empty set of possible worlds, and $\mathcal{A}$ is a σ–field over Ω. And as before, each possible world specifies a complete history, past, present, and future; but now agents belong to worlds, and as time goes by, their probabilities concerning their own and other agents' probabilities evolve along with their probabilities concerning the rest of the world—e.g., my probability for your probability for my probability for a Republican President in the year 2000.

The time index t takes values in some linearly ordered set T. Probability measures $pr_{i\omega t}$ represent ideally precise probabilistic states of mind of agents i in worlds ω at times t. The common prior P in Theorem 1.2 is such a measure; but in the present framework we need to spell out the commonality assumption, and in doing so we see that there might be different common priors in different worlds: perhaps the probability measures $pr_{i\omega t}$ are the same for all i, and so are $pr_{i\omega' t}$, but $pr_{i\omega t}(A) \neq pr_{i\omega' t}(A)$ for some A.

Now Def. B^t defines *Belief*—that is, certainty, full belief—as 100% probability, and Def. K^t replaces the old partition-based Def. K by a definition of knowledge simply as true full belief:

(Def. B^t) $$B_i^t A = \{\omega : pr_{i\omega t}(A) = 1\}$$

(Def. K^t) $$K_i^t A = A \cap B_i^t A$$

The old definitions of mutual and common knowledge are then adapted to the new definition of knowledge:

(Def. MK^t) $$M_0^t A = A$$
$$M_{n+1}^t A = \bigcap_{i=1}^{N} K_i^t M_n^t A$$

(Def. CK^t) $$\kappa^t A = \bigcap_{n=0}^{\infty} M_n^t A$$

In the generalized theorem, $t = 1$ and $t = 2$ are times at which agents have their old and new probabilities for some hypothesis $H \in \mathcal{A}$. In world ω those probabilities are $pr_{i\omega 1}(H)$ and $pr_{i\omega 2}(H)$—which we now write simply as $P_\omega(H)$ and $Q_\omega(H)$, respectively, with the subscript i understood, and with the work of the time subscripts 1 and 2 done by writing P and Q:

(Shorthand) $$P_\omega = pr_{i\omega 1}$$ $$Q_\omega = pr_{i\omega 2}$$

Thus, with $t = 2$, Def. B^t would be written:

(Def. B^2) $$B_i^2 A = \{\omega : Q_\omega(A) = 1\}$$

In the generalized theorem the proposition C^2 specifies $q_1, \ldots, q_N$ as the agents' new probabilities for H:

(Def. C^2) $$C^2 = \bigcap_{i=1}^{N} \{\omega : Q_\omega(H) = q_i\}$$

The crucial hypothesis of the generalized theorem is Goldstein's (1983) principle (G) that *old probabilities = old expectations of new probabilities.* Here the integrand $Q(A)$ is a random variable, a P_ω-measurable function of ω which takes real values $Q_{\omega'}(A)$ as ω' ranges over Ω :

(G) $$P_\omega(A) = \int_\Omega Q(A)\, dP_\omega$$

The second hypothesis says that whenever A is in $\mathcal{A}$, so are the N propositions saying—perhaps, falsely—that at time 2 the several agents are sure that A is true. This guarantees that $\mathcal{A}$ is closed under all the operations K_i^2, M_n^2, and κ^2.

To prove the theorem we use two lemmas. The first is the analog of the lemma (1.1) used to prove Aumann's theorem in Section 1.

(More Shorthand) $$B = B_i^2, \;\; \kappa = \kappa^2, \;\; C = C^2$$

Lemma 3.1 *While something is common knowledge, everyone is sure it is:* $\kappa \subseteq B\kappa$.

Proof. For the example, use Def. CK^2, MK^2, and K^2. ■

Lemma 3.2 *If (G) holds, then* $\int_{\Omega-\kappa C} Q(\kappa C)\, dP_\omega = 0$

Proof. $P_\omega(\kappa C) = \int_\Omega Q(\kappa C)\, dP_\omega = \int_{\kappa C} Q(\kappa C)\, dP_\omega + \int_{\Omega-\kappa C} Q(\kappa C)\, dP_\omega$ by (G). By lemma 3.1 the first term of this sum $= \int_{\kappa C} 1\, dP_\omega = P_\omega(\kappa C)$, so the second term $= 0$. ■

Theorem 3.3 (Aumann's "No Agreement" Theorem Generalized) *Hypotheses: (G) holds, B maps $\mathcal{A}$ into itself, P_ω is the same for all i, $P_\omega(\kappa C) > 0$. Conclusion: $P_\omega(H|\kappa C) = q_1 = \cdots = q_N$.*

Proof. $P_\omega(H|\kappa C)$

$= \dfrac{\int_\Omega Q(H\cap\kappa C)\,dP_\omega}{\int_\Omega Q(\kappa C)\,dP_\omega}$ by (G)

$= \dfrac{\int_{\kappa C} Q(H\cap\kappa C)\,dP_\omega + \int_{\Omega-\kappa C} Q(H\cap\kappa C)\,dP_\omega}{\int_{\kappa C} Q(\kappa C)\,P(d\omega) + \int_{\Omega-\kappa C} Q(\kappa C)\,dP_\omega}$

$= \dfrac{\int_{\kappa C} Q(H\cap\kappa C)\,dP_\omega}{\int_{\kappa C} Q(\kappa C)\,dP_\omega}$ by lemma 3.2, since $Q(A\cap\kappa C) \le Q(\kappa C)$

$= \dfrac{\int_{\kappa C} Q(H)\,dP_\omega}{\int_{\kappa C} 1\,dP_\omega}$ by lemma 3.1

$= \dfrac{\int_{\kappa C} q_i\,dP_\omega}{\int_{\kappa C}\,dP_\omega}$ by Def. $C^2 = q_i$, i.e., the same for all i. ∎

4 Remarks

4.1 Solo Epistemology

Aumann's theorem 1.2 was the start of what he would later (1995) call *interactive epistemology*. Here we use the term *solo epistemology* for the special case in which the number N of agents is 1. In this case Theorem 3.3 assumes the form of Corollary 4.1 below, with C and κC assuming especially simple forms:

(Solo C) $\qquad C = \{\omega : Q_\omega(H) = q\}$

(Solo κC) $\qquad \kappa C = C \cap BC \cap BBC \cap \cdots$

Corollary 4.1 *Hypotheses: (G) holds, B maps $\mathcal{A}$ into itself, $P_\omega(\kappa C) > 0$. Conclusion: $P_\omega(H|C\cap BC\cap BBC\cap\cdots) = q$.*

This says that in the presence of two more hypotheses, (G) implies a variant of van Fraassen's (1984) *reflection principle*:

(Reflection) $\qquad P_\omega(H|C) = q$

Goldstein (1983) and van Fraassen (1984) offer a scoring rule argument and a Dutch book argument for (G) and Reflection as diachronic coherence principles.

4.2 Tightness

The only properties of κC used the proof of Theorem 3.3 are $\kappa C \subseteq C$ (to get from line 5 to line 6) and $\kappa C \subseteq B\kappa C$ (to get from line 4 to line 5). In fact κC fits this pair of properties tightly, in the following sense.

Corollary 4.2 *κC is the largest $A \subseteq C$ for which $A \subseteq BA$.*

Proof. By Def. CK^2, κC satisfies both conditions on A. And by induction, any A satisfying both conditions is included in κC. (Basis: $A \subseteq M_0^2 A$ by Def. M_0^2. Induction step: If $A \subseteq M_n^2 A$ then as the second condition holds for all $i = 1, \ldots, N$, $A \subseteq M_{n+1}^2 A$. Then $A \subseteq \kappa C$ by Def. CK^2.) ■

4.3 Belief and Knowledge

In Aumann's framework, true belief, that is, 100% new probability for a truth, is definable (Def. K), but mere belief is indefinable. In the new framework belief is definable as in Def. B^t in Section 3, so that in the new framework Def. K^t defines knowledge as true belief.

The weakness of this notion of belief is indicated by the inventory (B0)–(B6) of conditions on mappings $M : \mathcal{A} \to \mathcal{A}$ that are satisfied ($\Uparrow$) and violated ($\Downarrow$) when M is belief, $M = B_i^t$, as in Def. B^t. But despite their apparent weakness, the four conditions it does satisfy suffice for the generalized "No Agreement" Theorem.

$\Uparrow$(B0) $(MA_1 \cap MA_2 \cap \ldots) \subseteq M(A_1 \cap A_2 \cap \ldots)$.
Distributivity

$\Uparrow$(B1) If $A_1 \subseteq A_2$ then $MA_1 \subseteq MA_2$
Deductive closure

$\Downarrow$(B2) $MA \subseteq A$
Infallibility

$\Downarrow$(B3) $MA \subseteq MMA$
Positive introspection

$\Downarrow$(B4) $\neg MA \subseteq M\neg MA$
Negative introspection

$\Uparrow$(B5) $M\emptyset = \emptyset$
Consistency

$\Uparrow$(B6) $M\Omega = \Omega$
Necessity

4.4 Demurrals

Since Plato's *Theaetetus* it has been a philosophical commonplace that true belief need not be knowledge. Then (Def. K^t) in Section 3 is suspect as an analysis of knowledge, as philosophers understand the matter. Nor are the credentials of Def. B^t as an analysis of what is normally called belief any better, for we may truly be said to believe propositions for which our probabilities fall slightly short of 100%. Nor can precise probability measures be taken seriously as general models of human judgmental states. Interactive epistemology is best understood as a branch of a theory of judgmental probability in which the notions of all-or-none belief and knowledge, and of precise judgmental probabilities, appear for the most part as limiting cases.

References

[1] Aumann, R.J. (1976), "Agreeing to Disagree", *Annals of Statistics* **4**, 1236–1239.

[2] Aumann, R.J. (1995), "Interactive Epistemology", Discussion Paper No. 67, Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem.

[3] Bacharach, M. (1985), "Some Extensions of a Claim of Aumann in an Axiomatic Model of Knowledge", *Journal of Economic Theory* **37**, 167–190.

[4] Billingsley, P. (1995), *Probability and Measure*, New York: John Wiley & Sons.

[5] Cave, A.K. (1983), "Learning to Agree", *Economics Letters* **12**, 147–152.

[6] Diaconis, P. And Zabell, S. (1982), "Updating Subjective Probabilities", *Journal of the American Statistical Association* **77**, 822–830.

[7] Goldstein, M. (1983), "The Prevision of Prevision", *Journal of the American Statistical Association* **78**, 817-819.

[8] Jeffrey, R.C. (1992), *Probability and the Art of Judgment*, Cambridge: Cambridge University Press.

[9] Samet, D. (1990), "Ignoring Ignorance and Agreeing to Disagree", *Journal of Economic Theory* **52**, 190–207.

[10] van Fraassen, B. (1984), "Belief and the Will", *Journal of Philosophy* **81**, 235-256.