Econometrica, Vol. 49, No. 1 (January, 1981)

AN AXIOMATIC CHARACTERIZATION OF COMMON KNOWLEDGE

BY PAUL MILGROM[1]

INTUITIVELY, AN EVENT A is common knowledge among a group of agents if each agent knows A, each knows that all know A, each knows that all know that all know A, etc. In this case, the "etc." encompasses an infinite sequence of conditions, each more stringent than the one before. In a recent paper, Aumann [1] formalized the idea of common knowledge in the following simple way.

Let (Ω, p) be a finite probability space and let $\mathscr{P}$ and $\mathscr{Q}$ be partitions of Ω representing the information of two agents. Let $\mathscr{R}$ be the meet of $\mathscr{P}$ and $\mathscr{Q}$, i.e. the finest common coarsening of the two partitions:

$$\mathscr{R} = \mathscr{P} \wedge \mathscr{Q}. \tag{1}$$

By the notation $\mathscr{P}(\omega)$ (resp. $\mathscr{Q}(\omega)$, $\mathscr{R}(\omega)$) is meant that element of $\mathscr{P}$ (resp. $\mathscr{Q}$, $\mathscr{R}$) which contains ω.

DEFINITION: An event A is common knowledge at $\omega (\omega \in \Omega)$ if $\mathscr{R}(\omega) \subset A$.

Aumann used this definition to state and prove a theorem asserting that two experts cannot "agree to disagree."

THEOREM 1: *Suppose that for some event B and some ω it is common knowledge at ω that $p(B|\mathscr{P}) = \alpha$ and $p(B|\mathscr{Q}) = \beta$. Then $\alpha = \beta$.*

Thus, when two experts have identical prior beliefs and each obtains some private information about B, if each knows the other's posterior beliefs, each knows that the other knows his beliefs, etc., then these posterior beliefs must be identical. Geanakoplos and Polemarchakis [2] extended this result by showing that if two experts simply communicate their posterior beliefs back and forth, then they "will be led to make revisions that converge, in finitely many steps, to the common, equilibrium posterior."

These results are set in a non-economic context,[2] but they suggest interesting economic questions. When traders exchange a risky security on the basis of private information, are they not "agreeing to disagree" in some formalizable sense? What sort of informational exchange leads to equilibrium beliefs in various trading mechanisms?

Recently, the idea of common knowledge has been extended to the case of many agents and applied to the analysis of a rational expectations trading model [3]. Formally, let partitions $\mathscr{P}_1, \ldots, \mathscr{P}_n$ represent the information of agents 1 through n, respectively, and let $\mathscr{R}$ be the meet of these n partitions:

$$\mathscr{R} = \mathscr{P}_1 \wedge \ldots \wedge \mathscr{P}_n. \tag{2}$$

Then common knowledge is defined just as above, with $\mathscr{R}$ interpreted according to (2).

In the study of trade among self-interested rational traders in uncertain environments, one cannot gain a good intuitive grasp of the standard results simply by recognizing that each trader knows that the others may be basing their decisions on private information. It is also necessary to recognize that each trader knows that the others know that each is using his information, etc.[3] Let us see how this point of view is used.

[1] I owe a debt of thanks to an anonymous referee whose comments led me to a clearer exposition of the applications of common knowledge.

[2] *The Economist* (December 29, 1979, p. 7) has offered a common knowledge style analysis of recent events in the international arena: "In the weird multibluff of nuclear arms, the Russians know the Americans believe the Russians think the Soviets could win a first strike nuclear war..."

[3] For an example which emphasizes this point, see [3, p. 12].

Suppose that there are n traders and l commodities. States of the world are points in the finite set $\Omega \subset \Theta \times X$. Each trader i is characterized by a strictly concave von Neumann-Morgenstern utility function $U_i : \mathbb{R}^l_+ \to \mathbb{R}$, a random endowment $e_i : \theta \to \mathbb{R}^l_+$, and a probability measure p_i on Ω. Note that although states of the world have the form $\omega = (\theta, x)$, traders' endowments depend only on θ. Thus, θ can be regarded as the "payoff-relevant" aspect of the environment. Of course, full or partial knowledge of x may convey useful information about θ.

A θ-contingent trade is defined to be an n-tuple $t = (t_1, \ldots, t_n)$ of functions $t_i : \Theta \to \mathbb{R}^l$. Each trader's information is represented by a partition in the usual way.

Now suppose the traders trade to an *ex ante* Pareto optimum before observing their private information. When the private information arrives, the marginal conditions will, in general, be disturbed, and markets might therefore be reopened. But can risk-averse traders really agree to a trade based solely on differences in information? A negative answer to this question is given in [**3**], and a slightly specialized version of that theorem is stated below.

THEOREM 2: *Suppose that the traders have identical prior beliefs* $(p_1 = \ldots = p_n)$ *and that the initial allocation* $(e_1, \ldots, e_n)$ *is Pareto optimal ex ante* (*before any information becomes available*) *relative to* θ*-contingent trades. Let t be a proposed* θ*-contingent trade ex post and suppose it is common knowledge at some* ω *that* (3)–(5) *hold* (*i.e., that the trade is feasible and acceptable to each trader*):

$$\sum_{i=1}^{n} t_i \leqslant 0, \tag{3}$$

$$\forall i \quad e_i + t_i \geqslant 0, \tag{4}$$

$$\forall i \quad E[U_i(e_i + t_i)|\mathcal{P}_i] \geqslant E[U_i(e_i)|\mathcal{P}_i]. \tag{5}$$

Then $t_1(\omega) = \ldots = t_n(\omega) = 0.$[4]

Additional applications of the common knowledge idea arise frequently. Wilson [**5**] defines an efficient allocation in a world of differential information in a way that can be stated succinctly using common knowledge: A contingent allocation x is *efficient* if there is no other allocation y such that it is common knowledge that all prefer y to x. Kobayashi [**4**] investigates the convergence of beliefs problem in a way that resembles the Geanakoplos and Polemarchakis work. Additional results using these ideas can be found in [**3**].

The applications cited above amply demonstrate the usefulness of the formal notion of common knowledge. Since the formal definition is regretably far-removed from the original intuitive idea, it may be useful to describe common knowledge in terms of its characteristic properties. Formally, I treat the problem of characterizing common knowledge as one of associating with each event A another event K_A with the interpretation

$$K_A = \{\omega \in \Omega | A \text{ is common knowledge at } \omega\}. \tag{6}$$

Consider the following four properties:

$$K_A \subset A, \tag{P1}$$

$$\forall \omega \in K_A, \quad \forall i, \quad \mathcal{P}_i(\omega) \subset K_A, \tag{P2}$$

$$B \subset A \Rightarrow K_B \subset K_A, \tag{P3}$$

$$[\forall i, \forall \omega \in A, \mathcal{P}_i(\omega) \subset A] \Rightarrow A = K_A. \tag{P4}$$

[4] The proof proceeds by showing that if the trade t is not null, then the trade t^* defined by $t_i^* = 1/2E[1_{\mathcal{R}(\omega)} t_i | \theta]$ is feasible and ex ante Pareto improving. Intuitively, gains from trade can arise in this setting only by systematically outguessing other traders. In a rational expectations equilibrium, some trader must realize that he cannot outguess his trading partners, so no trade takes place.

Condition (P1) asserts that A can be common knowledge only if A actually occurs. Conditon (P2) holds that if A is common knowledge, then every agent knows that A is common knowledge. Beginning with (P1) and applying (P2) repeatedly, one can infer that A is common knowledge *only if* A occurs, each agent knows A, each knows that all know A. etc. Condition (P3) holds that whenever B is common knowledge, any logical consequence of B is also common knowledge. Condition (P4) asserts that *public events* are common knowledge whenever they occur. The antecedent in (P4) defines a public event: it is an event which, if it occurs, will be known to every agent. I offer two examples of public events which arise in a trading context.

Suppose that trading is controlled by a Walrasian auctioneer who announces prices until some market clearing prices are found. We may imagine that at the close of trading, the auctioneer announces the equilibrium price vector p to the assembled traders. Then no trader can fail to know that "p is the equilibrium price vector" whenever such is the case. Similarly, no trader can fail to know that the net trades proposed by the various traders at the final round of the *tatonnement* process are feasible.

If trading takes place through some bargaining process, one may suppose that various proposals and counter-proposals will be made. At the close of bargaining when the traders sign the final documents, each trader must know that "this particular bargain is acceptable to everyone."

THEOREM 3: *There is a unique function K satisfying* (*P*1)–(*P*4) *and it is given by*

$$K_A = \{\omega | \mathcal{R}(\omega) \subset A\}. \tag{7}$$

PROOF: One can readily check that the function defined by (7) satisfies (P1)–(P4), thus proving existence. For uniqueness, take any event A. It follows from (P2) that $\{K_A, K_A^c\}$ is coarser than any of the partitions $\mathcal{P}_1, \ldots, \mathcal{P}_n$. Since $\mathcal{R} = \mathcal{P}_1 \wedge \ldots \wedge \mathcal{P}_n$ is the finest common coarsening, it follows that for all $\omega \in K_A$, $\mathcal{R}(\omega) \subset K_A$. Hence by (P1) $\mathcal{R}(\omega) \subset A$ for all $\omega \in K_A$, which is the meaning of the statement:

$$K_A \subset \{\omega | \mathcal{R}(\omega) \subset A\}. \tag{8}$$

Next suppose ω is such that $\mathcal{R}(\omega) \subset A$. Then by (P3), $K_{\mathcal{R}(\omega)} \subset K_A$. It is straightforward to check that $\mathcal{R}(\omega)$ is a public event, so using (P4) leads to $\mathcal{R}(\omega) \subset K_A$ and, in particular, to $\omega \in K_A$. This paragraph has shown that

$$\{\omega | \mathcal{R}(\omega) \subset A\} \subset K_A. \tag{9}$$

Taken together (8) and (9) establish uniqueness. *Q.E.D.*

COROLLARY: *Any definition of common knowledge consistent with* (6) *and* (*P*1)–(*P*4) *is equivalent to Aumann's definition.*

Notice that the proof given above implicitly defines *two* characterizations of common knowledge. The first part of the proof implies that K_A (as defined by (7)) is the *most inclusive* set consistent with (P1) and (P2). I argued earlier that (P1) and (P2) characterize the requirements of the intuitive definition of common knowledge.

The second part of the proof implies that K_A is the *least inclusive* set consistent with (P3) and (P4). So an event is common knowledge at ω if and only if it is the logical consequence of a public event that occurs at ω. Since the elements of $\mathcal{R}$ are all public events, this second characterization gives intuitive content to Aumann's formal definition.

Northwestern University

Manuscript received August, 1979; final revision received February, 1980.

REFERENCES

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

[2] GEANAKOPLOS, J., AND H. POLEMARCHAKIS: "We Can't Disagree Forever," Institute for Mathematical Studies in the Social Sciences, Technical Report No. 277, Stanford University, December, 1978.

[3] MILGROM, P., AND N. STOKEY: "Information, Trade, and Common Knowledge," Center for Math Studies, Discussion Paper No. 377R, Northwestern University, November, 1979.

[4] KOBAYASHI, T.: "Informational Efficiency of Competitive Prices: An Intertemporal Analysis," HBS 79-8, Harvard Business School, 1979.

[5] WILSON, R.: "Information, Efficiency, and the Core of an Economy," *Econometrica*, 46 (1978), 807–816.