Consensus By Identifying Extremists

Robin Hanson
California Institute of Technology
Pasadena, California

Abstract. Given a finite state space and common priors, common knowledge of the identity of an agent with the minimal (or maximal) expectation of a random variable implies "consensus", i.e., common knowledge of common expectations. This "extremist" statistic induces consensus when repeatedly announced, and yet, with n agents, requires at most $\log_2 n$ bits to broadcast.

Key words: consensus, common knowledge, information pooling, Bayesian learning

1. Introduction

Given a group of rational Bayesian agents, let a "knowledge property" be some statistic over the information held by individual agents. If these agents start from common prior beliefs, there are a number of knowledge properties where common knowledge of the property among the group is known to imply "consensus", i.e., common knowledge of common expectations regarding some random variable. And for most of these knowledge properties, it has also been shown that common knowledge of the property, and hence consensus, is eventually induced in finite worlds simply by repeatedly announcing the current value of the property.

For example, Nielsen et. al. (1990) showed that common knowledge of any "stochastically monotone" statistic, i.e., a sum of strictly monotonic functions of each agent's expectation, implies that these expectations are equal. And in a finite world, repeatedly announcing such a statistic eventually induces common expectations. Thus if a market clearing price is stochastically monotone in some expectation (several such models are in McKelvey & Page (1986)), consensus can be induced by repeatedly announcing such a market price.

Previously known consensus-inducing knowledge properties, however, share with this example the feature of being relatively direct and precise functions of or bounds on the numerical values of individual estimates. For example, announcing a stochastically monotone market price knowledge property would generically aggregate all private information with a single annoucement. Given substantial private information (a large state space and fine information partitions), doing so would require announcing a very precise price via a very long annoucement message. Thus this may not be a realistic model of actual information aggregation, in markets or elsewhere.

In contrast, this note presents a different sort of consensus-inducing knowledge property, one which directly concerns agent identity rather than agent estimate numbers. Because of this, given n agents at most $\log_2 n$ bits is needed to broadcast this property, regardless of the size of the state space or information partitions.

Specifically I show that, given common priors and finite information partitions, common knowledge of the identity of an agent with minimum (or maximum) expectation for a real-valued random variable implies consensus, i.e., common knowledge of common expectations, and that repeated announcement of this property implies eventual common knowledge of it.

This result thus raises hopes of creating economic models of convergence to consensus with more realistic message lengths. For example, a common value auction might reach convergence by repeatedly announcing the identity of the current high bidder. Or a community of experts might reach consensus via repeated media coverage of those with the most extreme views. This note, however, does not elaborate any such models.

2. Previous Work

Aumann (1976) first showed that common knowledge of all posterior beliefs regarding some event implies common posterior beliefs regarding that event. Sebenius and Geanakoplos (1983) then showed that if it is common knowledge among two agents that one agent's conditional expectation of some real-valued random variable is non-negative, and that the other agent's expectation is non-positive, then these two expectations must both be zero. (These results generalize trivially to any non-zero dividing point.)

McKelvey and Page (1986) allowed any number of agents, and showed that they need only have common knowledge of a single aggregate statistic of their posteriors, if this statistic is a sum of strictly monotonic functions of each posterior. Neilsen, Brandenburger, Geanakoplos, McKelvey, and Page (1990) extended this result from posteriors of an event to conditional expectations of any real-valued random variable. Neilsen (1995) generalized this result to statistics of multivariate random variables.

Geanakoplos and Polemarchakis (1982) showed with finite information partitions, the posteriors of two agents regarding some event become common knowledge after a finite number of steps of communicating these posteriors back and forth. Neilsen (1984) generalized this to information represented by sigma-algebras instead of partitions.

Sebenius and Geanakoplos (1983) showed that within a finite number of steps of communicating back and forth whether one agent's expectation is strictly positive, and whether the other agent's expectation is strictly negative, someone will deny one of these statements. McKelvey and Page (1986) showed that repeated announcement of their monotonic statistic induces common knowledge, and even tested this prediction in (McKelvey & Page, 1990) in a simple, though unfortunately flawed (Hanson,), experiment.

3. Notation

Following the notation of Nielsen et. al. (1990), let Ω be a finite set of states of the world ω, each with positive prior $P(\omega)$. Let each agent $i \in N$ start with the same common prior beliefs $P(\omega)$, and then receive private information according to a partition Π_i of Ω, and let $\Pi_i(\omega)$ denote the element of Π_i containing $\omega \in \Omega$. (And assume all i know all Π_i.)

Given any real-valued random variable $X(\omega)$, if the true state is ω^* then agent i will have a conditional expectation of X given by

$$X_i(\omega^*) \equiv E(X \mid \Pi_i(\omega^*)) \equiv \frac{\sum_{\omega \in \Pi_i(\omega^*)} X(\omega)P(\omega)}{\sum_{\omega \in \Pi_i(\omega^*)} P(\omega)} \tag{1}$$

For any true state ω, we say it is *common knowledge* among N that the true state is in $\Pi(\omega)$, where the partition $\Pi \equiv \wedge_{i \in N} \Pi_i$ is the meet (or finest common coarsening) of the partitions Π_i. We also call an event $A \subset \Omega$ common knowledge at ω if $\Pi(\omega) \subset A$, we call a predicate $\mathcal{P}$ common knowledge at ω if the event $\{\omega' \in \Omega \mid \mathcal{P}(\omega')\}$ is common knowledge, and we call a function f common knowledge at ω if for all $\omega' \in \Pi(\omega)$ we have $f(\omega') = f(\omega)$.

Finally, let a *knowledge property* be a function $F(\omega, \Pi_1(\omega), \Pi_2(\omega) \ldots)$ of the state ω and partition elements $\Pi_i(\omega)$. This includes any function of partition element probabilities $P(\Pi_i(\omega)) = \sum_{\omega' \in \Pi_i(\omega)} P(\omega')$ or expectations $E(X \mid \Pi_i(\omega))$. Such a knowledge property is *noisy* if the dependence on the first argument is non-trivial.

4. Common Knowledge of an Inequality

Sebenius and Geanakoplos (1983) prove their Proposition 2, a "no betting" theorem:

THEOREM 1. *For any two agents, if it is common knowledge that one's expectation is no less than some value, and that the other's expectation is no greater than this value, then it is common knowledge that these expectations are equal.*

This result is represented in Figure 1a. The axes are the expectations X_1 and X_2 of two agents regarding some bounded variable (with a maximum and minimum possible value), and so points in this graph are expectation pairs $\langle X_1, X_2 \rangle$. Theorem 1 says that if it is common knowledge that the expectation pair is somewhere in the shaded region, then it must be common knowledge that this pair is on the diagonal. This result is more general that Aumann's previous result of Aumann (1976), which required common-knowledge of the exact location of this pair $\langle X_1, X_2 \rangle$.

Using a small variation on Sebenius and Geanakoplos's proof of theorem 1, we can prove theorem 2 below, which requires only that it be common knowledge that the pair is in the larger region shown in Figure 1b. Since this region contains any region as in Figure 1a, this result subsumes the previous result, and is not implied by it.

THEOREM 2. *For any two agents, if it is common knowledge that one's expectation is no less than the other's, then it is common knowledge that these expectations are equal.*

Proof. By rearranging equation 1, we have

$$\sum_{\omega \in \Pi_i(\omega^*)} X_i(\omega)P(\omega) = \sum_{\omega \in \Pi_i(\omega^*)} X(\omega)P(\omega)$$

Summing this over the partition elements $\Pi_i(\omega^*) \subset \Pi(\omega^*)$, we get

$$\sum_{\omega \in \Pi(\omega^*)} X_i(\omega)P(\omega) = \sum_{\omega \in \Pi(\omega^*)} X(\omega)P(\omega) \tag{2}$$

Since the right hand side is independent of i, then for all $i, j \in N$ we must have

$$\sum_{\omega \in \Pi(\omega^*)} [X_i(\omega) - X_j(\omega)]P(\omega) = 0 \tag{3}$$

If it is common knowledge at ω^* that $X_i(\omega^*) \geq X_j(\omega^*)$, then this is true for all $\omega \in \Pi(\omega^*)$, and so equation 3 becomes a sum of non-negative terms set equal to zero, which can only be if each term in the sum is zero. Thus for all $\omega \in \Pi(\omega^*)$ we have $X_i(\omega) = X_j(\omega)$, and so this equality is common knowledge.

COROLLARY 3. *If it is common knowledge that no one has a lower (or greater) expectation than a particular agent, then it is common knowledge that all expectations are equal.*

Proof. If there is a j such that for all $i \in N$, it is common knowledge that $X_i(\omega) \geq X_j(\omega)$ (or that $X_i(\omega) \leq X_j(\omega)$), then by repeated application of theorem 2, for all $i \in N$ it is common knowledge that $X_i(\omega) = X_j(\omega)$. So is is common knowledge that they are all equal.

5. Common Knowledge Via Repeated Announcements

Most common knowledge results have a related result regarding repeated-announcements. For example, given our choice of a finite state space Ω, we can easily prove the following, using only a small variation on previous proofs (such as McKelvey & Page (1986)).

THEOREM 4. *If a knowledge property is repeatedly announced, then that property will become common knowledge in a finite number of steps.*

Proof. Recall that a (noisy) knowledge property at step t is a function F of the state ω and the partitions $\Pi_i^t(\omega)$ of agents at step t. By definition, an announcement regarding the value of property F at step

t results in this value being common knowledge at step $t+1$. That is, for all $\omega' \in \Pi^{t+1}(\omega)$,

$$F(\omega', \Pi_1^t(\omega'), \Pi_2^t(\omega'), \ldots) = F(\omega, \Pi_1^t(\omega), \Pi_2^t(\omega), \ldots)$$

Now every announcement either refines someone's information partition Π_i, or it informs no one. Given a finite state space Ω, only a finite number of such partition changes are possible, so there is a point after which announcements are uninformative. But if no partitions have changed from step t to $t+1$, then $\Pi_i^{t+1}(\omega) = \Pi_i^t(\omega)$, and so at step t this knowledge property F has become common knowledge at ω.

Consider the set of agents $\mathrm{argmin}_{i \in N} X_i(\omega)$, with minimal expectations of X. Any method of selecting one of these agents is a knowledge property, if it depends at most on the state ω and the information sets $\Pi_i(\omega)$. Thus we can easily combine corollary 3 and theorem 4 above to conclude the following.

COROLLARY 5. *Repeatedly announcing the identity of an agent who currently has the minimum (or maximum) expectation will, in a finite number of steps, result in common knowledge that these expectations are equal.*

6. Bounded Broadcasts

A well known result in information theory (see Cover & Thomas (1991)) is that the minimum expected number of (noiseless) bits B needed to communicate any variable $X(\omega)$ to an agent known to have prior $P(\omega)$ is $B = -\sum_x p(x) \log_2 p(x)$, where $p(x) \equiv P(\{\omega : X(\omega) = x\})$, and where we adopt the convention that $0 log 0 = 0$. It is also well known that if there are at most n possible values of X, so that $|\{x : p(x) > 0\}| \leq n$, then $B \leq \log_2(n)$.

If a knowledge property F is *broadcast* at stage t, so that each agent i receives the same message, then the common knowledge prior $P(\omega)/P(\Pi^t(\omega))$ is the natural choice for encoding F. Thus the expected number of bits $B^t(\omega^*)$ required to announce F at step t in state ω^* is $B^t(\omega^*) = -\sum_f p(f) \log_2 p(f)$, where

$$p(f) \equiv \frac{P(\{\omega \in \Pi(\omega^*) : F(\omega, \Pi_1^t(\omega), \Pi_2^t(\omega), \ldots) = f\})}{P(\Pi^t(\omega^*))}.$$

Since there are never more possible extremists than there are agents, we can conclude the following.

COROLLARY 6. *Announcing the identity of an agent with the minimum (or maximum) expectation requires an expected message length of no more than* $\log_2 |N|$ *bits, independent of the information structure* $\Omega, P,$ *and* Π_i.

In contrast, no similar message length bounds are known to hold for the other knowledge properties known to induce consensus by repeated announcement. For example, consider McKelvey and Page's (1986) "stochastically monotone" (non-noisy) knowledge property, a sum of monotonic functions of each agent's expectation X_i.

With a generic prior and random variable, such a knowledge property F, when applied to the initial information partitions Π_i, can easily give a different value $f = F(\omega, \Pi_1(\omega), \Pi_2(\omega), \ldots)$ in each element of the join (or coarsest common refinement) $\widehat{\Pi} \equiv \vee_{i \in N} \Pi_i$, and hence induce full information pooling on the first announcement. This is the typical behavior, for example, of the class of environments used in the McKelvey and Page (1990) experiments, and is closely related to the common phenomena of prices inducing complete information pooling in rational expectations models.

Since inducing full information pooling in one announcement requires communicating all private information in one announcement, the message length required to broadcast this announcement grows without bound as the total amount of private information grows without bound. Specifically, it would require an expected

$$\sum_{\pi \in \widehat{\Pi}} P(\pi|\Pi(\omega)) log P(\pi|\Pi(\omega))$$

bits to broadcast all private information.

7. Conclusion

When priors beliefs are common and information partitions are finite, common knowledge of the identity of an agent with extreme expectations implies consensus, i.e., common expectations for perfectly rational agents. Consensus can thus be induced simply by repeatedly announcing the identity of an extremist.

In contrast to the other consensus-inducing knowledge properties, only a bounded, and moderate, message length is required to broadcast an extremist identity. This property thus offers the possibility of models of convergence to consensus with more realistic announcement-communication bandwidths.

Acknowledgements

I am grateful to Richard McKelvey, John Ledyard, and several anonymous referees for helpful comments on earlier drafts of this paper.

References

Aumann, R.: 1976, 'Agreeing to disagree', *The Annals of Statistics* **4**(6), 1236–1239.
Cover, T. and Thomas, J.: 1991, *Elements of Information Theory*, John Wiley & Sons, New York.
Geanakoplos, J. and Polemaarchakis, H.: 1982, 'We can't disagree forever', *Journal of Economic Theory* **28**, 192–200.
Hanson, R.: 1996, 'Correction to McKelvey and Page, 'Public and private information: An experimental study of information pooling'', *Econometrica* **64**(5), 1223–1224.
McKelvey R., and Page, T.: 1986, 'Common knowledge, consensus, and aggregate information', *Econometrica* **54**(1), 109–127.
McKelvey, R. and Page, T.: 1990, 'Public and private information: An experimental study of information pooling', *Econometrica* **58**(6), 1321–1339.
Nielsen, L.T., Brandenburger, A., Geanakoplos, J., McKelvey, R., and Page, T.: 1990, 'Common knowledge of an aggregate of expectations', *Econometrica* **58**(5), 1235–1239.
Nielsen, L.T.: 1984, 'Common knowledge, communication, and convergence of beliefs', *Mathematical Social Sciences* **8**, 1–14.
Nielsen, L.T.: 1995, 'Common knowledge of a multivariate aggregate statistic', *International Economic Review*, **36**(1).
Sebenius, J., and Geanakoplos, J.: 1983, 'Don't bet on it: Contingent agreements with asymmetric information', *Journal of the American Statistical Association* **78**, 424–426.

Address for correspondence: H&SS 228-77, Caltech, Pasadena, CA 91125. *hanson@hss.caltech.edu*