When Rational Disagreement is Impossible

KEITH LEHRER

CENTER FOR ADVANCED STUDY IN THE BEHAVIORAL SCIENCES
AND THE UNIVERSITY OF ARIZONA

It is natural to assume that a group sharing the same information and respecting each other's opinions may reasonably disagree. I shall prove, on the contrary, that if the members of such a group search for truth and accept the mathematical implications of their state, then they must converge toward consensus. Disagreement is demonstrably irrational.[1]

The proof I shall offer will demonstrate that a group of people who give some weight to the probability assignments of others will converge toward a consensual probability assignment. The probability assignment, p_i^0, of an individual i at stage 0 may, for my purposes, be thought of either as determining reasonable acceptance[2] or as constituting degrees of belief.[3] The principal thesis of my paper is best sustained, however, by interpreting the probability assignment of an individual as his estimate of the chances a proposition has of being true.

My first assumption is that each member of the group has at least some small degree of positive respect for the probability assignments of other members of the group. This degree of respect is formulated as a weight which he gives to the probability assignment of others in the group. Thus, w_{ij} is the weight that person i gives to the probability assignment of person j. A person assigns a positive weight to each probability assignment, including his own, and the set of weights, w_{i1}, w_{i2}, and so forth to w_{in}, which a person i assigns to the n probability functions, sum to 1. I call this the *respect* assumption. It amounts to each person giving some positive weight to the opinion of others. Actually, a much weaker

NOÛS 10 (1976)

assumption will suffice for the proof we shall offer. Even if some members of the group give no positive weight to other members of the group, convergence of probability assignments will be obtained if there is a vector of positive respect from each member to each other member. There is a vector of positive respect from individual i to j if there is a sequence of members, the first of which is i and the last of which is j, such that each member of the sequence gives positive weight to the probability assignment of the next member of the sequence. I shall construe these weights as second level probabilities that a person attaches to the various first level probability assignments of members of the group.[4]

Imagine a situation in which the members of the group alluded to are a group of scientific experts who satisfy the respect condition when the probability assignments of the experts are on a partition in the domain of their expertise. Imagine, moreover, that the assignment of weights is subsequent to a complete exchange of information, concerning both experimental data and theoretical extrapolation, relevant to the subject matter described. This I shall refer to as the *initial state.* Our second assumption is that, as long as no new information, experimental or theoretical, is introduced, the weights the members of the group give to the probability assignments of others remain constant. I call this the *constancy* assumption. The assumption is justified because the weights each person gives to the probability assignments of persons in the group represents his estimate of the chances each person has of being correct in his probability assignment. With these assumptions, mathematical reasoning yields the result that the application of weights to probabilities will produce convergent probability assignments.

We may now assume a principle of rationality, that a person in the initial state must, if he is rational, aggregate the initial probability assignments using the weights he assigns to obtain a new probability assignment. Thus, a person i in state 0 shifts to state 1 by the following method: $p_i^1 = w_{i1}p_1^0 + w_{i2}p_2^0 + \ldots + w_{in}p_n^0$. The probability assignment of i in state 1 is the weighted average of the probability assignments in state 0. All members of the group having shifted from state 0 to state 1, the same procedure is applied to the probability assignments of members of the group in state 1 to shift to state 2. The method of aggregation by which a person i shifts from state m to state m +

1 is, therefore, as follows: $p_i^{m+1} = w_{i1}p_1^m + w_{i1}p_2^m + \ldots + w_{in}p_n^m$.

From this method of aggregation our result follows directly from theorems concerning Markov chains. The weights of people in the group are represented by the following matrix:

$$A = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ w_{n1} & w_{n2} & & w_{nn} \end{pmatrix} \qquad p^kE = \begin{pmatrix} p_1^kE \\ p_2^kE \\ \cdot \\ \cdot \\ \cdot \\ p_n^kE \end{pmatrix}$$

The matrix is square with n rows and n columns, the rows representing the set of weights a person assigns to the probability assignments of members of the group, and the columns representing the set of weights members of the group assign to a probability assignment of a single member at each stage. This matrix we call A. For proposition E, we let p^kE be a column vector with components p_1^kE, p_2^kE, $\ldots$ p_n^kE, that is, p^kE is the vector of probabilities of members of the group at stage k. By taking p^0E as a column vector multiplied by matrix A, we obtain $Ap^0E = p^1E$, and, more generally, if we take p^kE as a column vector multiplied by A, we obtain $Ap^kE = p^{k+1}E$.[5]

It is a general property of such matrices, regular Markov transition matrices, that iterated multiplication yields convergence, that is, as k increases, p^kE converges toward a set of equal probabilities. This result is a consequence of two general theorems. First, $p^kE = A^kp^0E$, that is, the original probability column vector, p^0E, multiplied by the matrix A to the power k, A^k.[6] Second, as the power k of A^k increases, we obtain convergence to a matrix in which all rows of the matrix are the same.[7] This is equivalent to a condition in which each member of the group assigns the same weights to all other members of the group. Aggregation under that condition yields the same probability assignment for each member of the group. There is a social probability, pE, toward which p_i^kE converges as k increases for every i and for every E. When the assumptions stated are satisfied, rational aggregation yields consensus. The personal probability assignment of each individual member of the group converges toward a common consensual probability assignment. I call this the *probability consensus theorem.*

Moreover, the respect assumption used in the proof may be weakened. What is required is that the matrix be regular. The matrix is regular if some power of the matrix is such that all components of the resulting matrix are positive. In more intuitive terms, this occurs when there is no subgroup of members of the group who give positive weight to the probability of assignments of each other but give 0 weight to all other members of the group. We obtain a regular matrix even if some members have no direct positive respect for others in the group, provided positive respect is *communicated* indirectly from each member of the group to each other member by some vector of positive respect. As we noted earlier, there is a positive vector from i to j when there is a sequence of members, the first member being i and the last member being j, such that each member in the sequence assigns some positive weight to the probability assignment of the next member.[8]

I shall now offer some defense of the assumptions that lead to the theorem. First, the respect assumption, weakened as indicated, may be taken as a condition of a *community* of experts. If some members of a group respect each other, give positive weight to the probability assignments of each other, but give no weight to the probability assignments of others, then they form a separate and distinct community. Only when each member of a group communicates respect for each other member, either directly or through a chain, does a community of inquiry exist.

The constancy condition is sustained by the assumption that members of the community are genuine truth seekers who, after a complete exchange of information, acquire no new information. The weights an individual assigns to others represent his estimate of how reliable members of the group are as indicators of truth. The constancy assumption amounts to the requirement that a person who forms an estimate of the reliability of others as indicators of truth apply that estimate consistently until he obtains new information. By shifting from the initial state 0 to a state k in which consensus is approached, members of the community are simply applying their estimates in this way.

Perhaps the most natural objection to this argument concerns the iterated procedure of aggregation. It may be conceded that the first shift, from state 0 to state 1, is required by rationality, but objected that this shift takes account of the probability assignments of others and no further aggregation is appropriate. The shift from state 0 to state 1 is assumed to be motivated by the search for truth and the estimates we form of the chances members of the group have of obtaining truth. We may assume, moreover, that the weights people assign to each other are known in the initial state. Thus, the result of a shift from state 0 to state 1 does not supply new information; the result of that shift is a mathematical consequence of the information of members in the initial state. Hence, refusing to shift from state 1 to state 2 is equivalent to assigning a weight of 0 to other members of the group at this stage. This amounts to the assumption that there is no chance that one is mistaken and no chance that others in the group with whom one disagrees are correct. In short, the only alternative to the iterated aggregation converging toward a consensual probability assignment is individual dogmatism at some stage.

Actual disagreement among experts must result either from an incomplete exchange of information, individual dogmatism, or a failure to grasp the mathematical implications of their initial stage. What is impossible is that the members of some community of inquiry should grasp the mathematical implications of their initial state and yet disagree. Perhaps the most important philosophical application of the probability consensus theorem concerns the ideal case. It shows us that a community of mathematically perfected truth seekers would reach consensus on the basis of shared information. Such a consensual probability assignment of members of an ideal community of inquirers may be conceived of as an intersubjectively rational probability assignment the individual truth seekers strives to obtain.

REFERENCES

[1] Jeffrey, Richard C., *The Logic of Decision* (New York: McGraw Hill, 1965).
[2] Kemeny, John G. and Snell, J. Laurie, *Finite Markov Chains* (New York: Van Nostrand, 1960).
[3] Thompson, Gerald L., *Introduction to Finite Mathematics* (Englewood Cliffs: Prentice-Hall, 1956).
[4] Lehrer, Keith, "Induction and Conceptual Change," *Synthese* 23 (1971): 206-25.
[5] ——, "Truth, Evidence, and Inference," *American Philosophical Quarterly* 11 (1974): 79-92.
[6] Savage, Leonard J., *The Foundations of Statistics* (New York: John Wiley and Sons, 1954).

NOTES

[1] Kit Fine suggested the basic theorem of the paper to me, and Gerald Kramer showed me how the theorem could be proved by using the Markov transition matrix.

[2] In [4] and [5], I show how probabilities determine what it is reasonable for a person to accept as evidence and as hypotheses inferred from evidence.

[3] For two major works advocating this position, see Savage [6] and Jeffrey [1].

[4] When we interpret these weights as second level probabilities, these probabilities are normalized to sum to one and, therefore, are comparative probabilities only. Thus, if I assign a weight of 4/10 to one person and 1/10 to another, this means only that the former is four times as likely to be correct as the latter in my opinion. We obtain the normalized probabilities from a set of ratios, r_1, r_2, and so forth to r_n, representing the probabilities person i assigns before normalization to each member of the group by letting r equal the sum of these ratios and letting the weight w_{ij} which person i assigns to j be r_j/r.

[5] This method is restricted so that E is a member of a partition P on which the original probability assignment is made. The probability, p^kS, of a statement S that is not a member of the partition is computed from the probabilities of members of the partition by summation rather than by aggregation. Multiplying a matrix A by column vector p^kE, Ap^kE, is accomplished by multiplying the top member of the column by the left member of a row, the second from the top by the second from the left in that row, and so forth, and taking the sum of those products as the component in that row in the new column. Hence, the second from the top member of the new column, $p^{k+1}E$, would be $p_1^kEw_{21} + p_2^kEw_{22} + \ldots + p_n^kEw_{2n}$. An elementary introduction to the mathematics of matrices and a statement (though not proof) of the relevant theorem is in Kemeny, *et al.* [3]: 178-205, 217-36, esp 220.

[6] To multiply a square matrix times itself, each column of the matrix is multiplied by the matrix in the manner described in note 4 to obtain the appropriate column of the new matrix. The theorem simply tells us that multiplying a square matrix with n rows and columns times a column with n rows or members is the same thing as multiplying the matrix times itself n times and multiplying the result by the column vector. Thus, multiplying the matrix times the initial column vector twice, that is AAp^0E, is equal to A^2p^0E.

[7] The theorem is given and proved in Kemeny and Snell [2]: 70-1.

[8] Cf. Kemeny and Snell [2]: 36. The relation between communication and regularity is formulated there. A regular transition matrix is one that contains all positive entries for some power of the matrix, and a matrix with communication of the sort described provides regularity.