MASTERCLASS

Rational Consensus in Science and Society

(Lehrer and Wagner, 1981)

Department of Philosophy, Logic, and Scientific Method

The London School of Economics and Political Science

20 October 2009

Second Session (14.30 – 16.00)

THE AXIOMATICS OF AGGREGATION

Carl Wagner
Department of Mathematics
The University of Tennessee

1. Why weighted *arithmetic* means?

Answer: They furnish the simplest *allocation aggregation methods.*

The Allocation Aggregation Problem: Suppose that each of n individuals is asked to assess the most appropriate values of some set of numerical decision variables $x_1,\ldots,x_m$. Values are constrained to be nonnegative, and to sum to some fixed positive real number s. How should their possibly differing individual assessments be aggregated into a single group assessment?

● Record their individual assessments in an n x m matrix $A = (a_{ij})$ where a_{ij} denotes the value assigned by individual i to variable x_j. Any such matrix is called an *s-allocation matrix.* If n = 1, it is called an *s-allocation row vector.*

Reformulation of the allocation aggregation problem: Given an s-allocation matrix $A = (a_{ij})$, produce an s-allocation row vector $a = (a_1,\ldots,a_m)$ that incorporates the assessments recorded in A in some reasonable way.

Two possible approaches, modeled on paradigms from social choice theory:

1. Single profile (following Bergson-Samuelson)—more on this in this evening's seminar.

2. Multi-profile (following Black and Arrow), which we pursue here.

● $\mathcal{A}(n,m;s)$ = the set of all n x m s-allocation matrices.

● $\mathcal{A}(m;s)$ = the set of all m-dimensional s-allocation row vectors.

● Any function F: $\mathscr{A}(n,m;s) \rightarrow \mathscr{A}(m;s)$ is called an *allocation aggregation method* (AAM). Each AAM F furnishes a method, applicable to every conceivable s-allocation matrix A, of reconciling the possibly different opinions recorded in A in the form of the group assignment $F(A) = a = (a_1,\ldots,a_m)$.

● Notation

A_j denotes the j^{th} column of matrix A.

a_j denotes the j^{th} entry of row vector a.

$\underline{\mathbf{c}}$ denotes the n x 1 column vector with all entries equal to c.

● Aggregation Axioms

Irrelevance of Alternatives (IA): For each $j = 1,\ldots,m$, and for all A, B in $\mathcal{A}(n,m;s)$, $A_j = B_j \Rightarrow F(A)_j = F(B)_j$.

Remark. IA is clearly equivalent to the existence of functions $f_j : [0,s]^n \to [0,s]$, $j = 1,\ldots,m$, such that for all A in $\mathcal{A}(n,m;s)$,

$$F(A)_j = f_j(A_j) \text{ and } \sum_{1 \le j \le m} f_j(A_j) = s.$$

Zero Preservation (ZP): For each $j = 1,\ldots,m$, and for all A in $\mathcal{A}(n,m;s)$, $A_j = \underline{\mathbf{0}} \Rightarrow F(A)_j = 0$, i.e., $f_j(\underline{\mathbf{0}}) = 0$ for each $j = 1,\ldots,m$.

Theorem 1.1. (L & W 1981) If $m \ge 3$, an AAM F satisfies IA and ZP if and only if there exists a *single* sequence $w_1,\ldots,w_n$ of weights, nonnegative and summing to 1, such that for all $A = (a_{ij})$ in $\mathcal{A}(n,m;s)$ and $j = 1,\ldots,m$,

$$F(A)_j = f_j(A_j) = w_1 a_{1j} + w_2 a_{2j} + \cdots + w_n a_{nj}.$$

Note that IA and ZP allow for *dictatorial aggregation* (for some fixed d in {1,…,n}, $w_d = 1$ and $w_i = 0$ for $i \neq d$).

Theorem 1.2. (Aczel, Ng, Wagner 1984) If $m \geq 3$, an AAM F satisfies IA if and only if there exist "weights" $w_1,\ldots,w_n$ ε [-1,1] and real numbers $\beta_1,\ldots,\beta_m$ ε [0,s] satisfying

$$(1) \quad -s\Sigma^{-} w_i \leq \beta_j \leq s(1 - \Sigma^{+} w_i)\ ,\ j = 1,\ldots,m,$$

where Σ^{-} indicates the sum of the negative weights and Σ^{+} the sum of the positive weights, and

$$(2) \quad \sum_{1 \leq j \leq m} \beta_j = (1 - \sum_{1 \leq i \leq n} w_i)\, s,$$

such that, for all A = (a_{ij}) in $\mathscr{A}$(n,m;s),

$$(3) \quad F(A)_j = f_j(A_j) = w_1a_{1j} + w_2a_{2j} + \cdots + w_na_{nj} + \beta_j.$$

● If $w_i \equiv 0$, aggregation is *imposed*.

● The weights w_i may be *negative*, subject to conditions (1) and (2). In particular, condition (1) implies that $\Sigma\, |w_i| \leq 1$, and hence that $\Sigma\, w_i \leq 1$.

● If $\Sigma\, w_i = 1$, then $\beta_j = 0$ for all j, and each $w_i \geq 0$. So an AAM F satisfying IA differs from simple weighted arithmetic averaging if and only if $\Sigma\, w_i < 1$. In such a case the formula for $F(A)_j = f_j(A_j)$ may be recast in the form

(4) $F(A)_j = f_j(A_j) = \Sigma_i\, w_i\,(a_{ij} - \sigma_j\,) + \sigma_j$

$= \Sigma_i\, w_i\, a_{ij} + [1 - \Sigma_i\, w_i]\, \sigma_j$,

where $\sigma_j = \beta_j \,/\, (1 - \Sigma_i\, w_i) \geq 0$. Here, $\Sigma_j\, \sigma_j = s$.

Example with negative weights: In (4), let $w_1 = \cdots = w_{n-1} = 0$, $w_n = -1/(m – 1)$, and $\sigma_1 = \cdots = \sigma_m = s/m$ => $f_j(A_j) = (s – a_{nj})/(m – 1)$.

● *A necessary and sufficient condition for all weights* w_i *to be nonnegative:* For all vectors X, Y ε $[0,s]^n$, and for each j = 1,…,m,

$$X \geq Y \;\; \Rightarrow \;\; f_j(X) \geq f_j(Y). \quad (\textit{weak dominance})$$

Exercise: Determine the consequences of requiring *strong dominance*, i.e.,

$$X \geq Y \;\; \Rightarrow \;\; f_j(X) \geq f_k(Y) \;\; \text{for all } j,k \in \{1,\ldots,m\}$$

● *The case of infinitely many decision variables:*

The above theorems also hold, with the very same proofs, when there are denumerably infinitely many decision variables x_1, x_2, …

But in the infinite case, IA *forces all weights* w_i *to be nonnegative* (exercise).

2. Aggregating Probability Measures.

● If Ω is a finite or denumerably infinite set, a function p: $\Omega \to [0,1]$ is called a *probability mass function* if $\sum_{\omega \in \Omega} p(\omega) = 1$.

The above theorems apply to the aggregation of probability mass functions when $|\Omega| \geq 3$.

Each probability mass function p on a countable set Ω gives rise to a set function P: $2^{\Omega} \to [0,1]$ defined for all subsets E of Ω by

$$P(E) := \sum_{\omega \in E} p(\omega).$$

P is a *discrete probability measure*.

● *Aggregating arbitrary probability measures*:

● If Ω is a set of any cardinality, a family **A** of subsets (called *events*) of Ω is called a *sigma algebra* if

(i) $\Omega \varepsilon \mathbf{A}$,

(ii) $E \varepsilon \mathbf{A} \Rightarrow E^c \varepsilon \mathbf{A}$, and

(iii) $E_1, E_2 \ldots \varepsilon \mathbf{A} \Rightarrow E_1 \cup E_2 \cup \cdots \varepsilon \mathbf{A}$.

● A function $P: \mathbf{A} \to [0,1]$ is called a *probability measure on* **A** if

(i) $P(\Omega) = 1$, and

(ii) If $E_1, E_2, \ldots$ is a sequence of pairwise disjoint events, then

$$P(E_1 \cup E_2 \cup \cdots) = P(E_1) + P(E_2) + \cdots \quad .$$

● $\Pi_{\mathbf{A}}$: = the set of all probability measures on the sigma algebra **A**.

● Any $F: (\Pi_{\mathbf{A}})^n \to \Pi_{\mathbf{A}}$ is a *probability aggregation method* (PAM).

Here, IA takes the form : For each E ε **A** (except the empty set and Ω) there exists a function $f_E : [0,1]^n \rightarrow [0,1]$ such that

$$F(P_1,\ldots, P_n)(E) = f_E(P_1(E),\ldots, P_n(E)), \text{ and}$$

ZP dictates that $f_E(0,\ldots,0) = 0$. The condition $m \geq 3$ is replaced by the requirement that **A** be *tertiary,* i.e., that there exist at least three nonempty, pairwise disjoint events in **A**.

Then IA and ZP characterize the PAMs

$$F(P_1,\ldots, P_n) = w_1 P_1 + \cdots + w_n P_n,$$

and IA alone characterizes the PAMs

$$F(P_1,\ldots, P_n) = w_1 P_1 + \cdots + w_n P_n + (1-\Sigma w_i)Q,$$

where Q is a probability measure on **A**.

3. Remarks on Irrelevance of Alternatives.

● Kevin McConway (Marginalization and linear opinion pools, *J.Amer.Statist. Assoc.* 76 (1981), 410-414) proved that IA is equivalent to a certain *marginalization property* of probability aggregation:

● Let $\mathcal{S}(\Omega)$ denote the set of all sigma algebras on Ω. For each **A** ε $\mathcal{S}(\Omega)$, let $F_{\mathbf{A}} : (\Pi_{\mathbf{A}})^n \rightarrow \Pi_{\mathbf{A}}$.

● A probability aggregation method (in the sense of McConway) is a family $\{ F_{\mathbf{A}}: \mathbf{A} \in \mathcal{S}(\Omega)\}$ of such mappings.

● Given **A** and **B** in $\mathcal{S}(\Omega)$, where **B** is a sub-sigma algebra of **A**, and P a probability measure on **A**, let $P_{(\mathbf{B})}$ denote the marginalization (i.e., the restriction) of P to **B**.

● The family $\{ F_{\mathbf{A}}: \mathbf{A} \in \mathcal{S}(\Omega)\}$ has the *marginalization property* (MP) iff

$$F_{\mathbf{B}}(P_{1(\mathbf{B})},\ldots,P_{n(\mathbf{B})}) = (F_{\mathbf{A}}(P_1,\ldots,P_n))_{(\mathbf{B})}$$

for all $(P_1,\ldots, P_n) \in (\Pi_{\mathbf{A}})^n$.

MP ⇔ marginalization commutes with aggregation.

Theorem 3.1 (McConway). The family $\{ F_{\mathbf{A}}: \mathbf{A} \in \mathcal{S}(\Omega)\}$ has the MP iff, for each nonempty, proper subset E of Ω, there exists a function $f_E: [0,1]^n \rightarrow [0,1]$ such that, for all $\mathbf{A} \in \mathcal{S}(\Omega)$, all $(P_1,\ldots, P_n) \in (\Pi_{\mathbf{A}})^n$, and all $E \in \mathbf{A}$,

(5) $F_{\mathbf{A}}(P_1,\ldots,P_n)(E) = f_E(P_1(E),\ldots,P_n(E))$.

Note that (5), which McConway calls the *strong setwise functionality property* (SSFP) implies IA for each $F_{\mathbf{A}}$, where $\mathbf{A} \in \mathcal{S}(\Omega)$. Morever, if **A**, **B** ε $\mathscr{S}(\Omega)$, E ε **A** ∩ **B**, $(P_1,\ldots, P_n)$ ε $(\Pi_{\mathbf{A}})^n$, $(Q_1,\ldots, Q_n)$ ε $(\Pi_{\mathbf{B}})^n$, and $P_i(E) = Q_i(E)$, i =1,…,n, then

$$F_{\mathbf{A}}(P_1,\ldots,P_n)(E) = F_{\mathbf{B}}(Q_1,\ldots,Q_n)(E).$$

4. Weighted Arithmetic Aggregation and Conditionalization.

Let F: $(\Pi_{\mathbf{A}})^n \rightarrow \Pi_{\mathbf{A}}$ be given by the formula

$$F(P_1,\ldots, P_n) = w_1P_1 + \cdots + w_nP_n.$$

If E ε **A**, then, in general,

$$F(P_1,\ldots, P_n)(\cdot|E) \neq w_1P_1(\cdot|E) + \cdots + w_nP_n(\cdot|E).$$

"Weighted arithmetic aggregation does not commute with conditionalization."

In fact, Dalkey (1975) showed that such commutativity holds iff aggregation is dictatorial. Is this a problem?

Note that $F(P_1,\ldots, P_n)(A|E) :=$

$$F(P_1,\ldots, P_n)(A\cap E) / F(P_1,\ldots, P_n)(E)$$

$$= \Sigma\, w_iP_i(A\cap E) / \Sigma\, w_iP_i(E)$$

$$= u_1P_1(A|E) + \cdots + u_nP_n(A|E), \quad \text{where}$$

$$u_i = w_iP_i(E) / \Sigma\, w_iP_i(E).$$

McConway: No problem if aggregation "applies only to distributions conditional on a fixed amount of knowledge."

But

$$(P_1,\ldots, P_n) \in (\Pi_{\mathbf{A}})^n \Rightarrow (P_1(\cdot|E),\ldots, P_n(\cdot|E)) \in (\Pi_{\mathbf{A}})^n$$

and the domain of the PAM F is assumed to be *all of* $(\Pi_{\mathbf{A}})^n$.

● Commutativity of aggregation and conditionalization can be achieved if

i. a weaker form of IA is adopted; and

ii. the probability measures are discrete, and aggregated via their associated mass functions.

Theorem 4.1. If $\Omega = \{\omega_1,\omega_2,\ldots\}$, M_Ω = the set of all probability mass functions on Ω. For all $(p_1,\ldots,p_n) \varepsilon (M_\Omega)^n$, and each j = 1,2,…, let

$F(p_1,\ldots,p_n)(\omega_j) :=$

$$\Pi_{1 \le i \le n}\ p_i(\omega_j)^{w(i)} / \ \Sigma_j (\Pi_{1 \le i \le n}\ p_i(\omega_j)^{w(i)}),$$

the *normalized weighted geometric mean of* $p_1(\omega_j),\ldots, p_n(\omega_j)$. Then the PAM F commutes with conditionalization (and also with Jeffrey conditionalization, parameterized, following H. Field, in terms of Bayes factors).

● Theorem 4.1 also holds for probability measures P_i on a sigma algebra **A** for which there exists a measure μ on **A** and μ-measurable "density functions" φ_i on Ω such that for all E ε **A**,

$$P_i(E) = \int_E \varphi_i \, d\mu.$$

Here, $F(\varphi_1,\ldots, \varphi_n) = \prod_i \varphi_i^{w(i)} / \int \prod_i \varphi_i^{w(i)} \, d\mu.$

5. Allocation Aggregation with a Finite Valuation Domain (Bradley and Wagner)

Suppose that the values assigned to the variables must lie in the finite subset V of [0,s], where
(i) 0 ε V.
(ii) x ε V => s – x ε V.
(iii) x, y ε V and x + y ≤ s => x + y ε V.

Theorem 5.1. If m ≥ 3, an AAM

F: $\mathscr{A}$(n,m;s,V) $\rightarrow$ $\mathscr{A}$(m;s,V) satisfies IA and Z if and only if it is dictatorial.