Strategic manipulation in Bayesian dialogues

Christina Pawlowitsch

christina.pawlowitsch@u-paris2.fr

Université Panthéon-Assas, Paris II

April 2021

Motivation

As individuals participating in society (consumers, investors, voters, citizens) we are frequently faced with situation in which we have to attribute probabilities to certain events — without being able to derive those probabilities from a clearly defined underlying mathematical model, such as the throw of a dice.

Instead, we have to exploit subjective information that we acquire about the state of the world.

Still we are rational. We want to come up with these subjective probabilities not in an arbitrary way. Rather we want to exploit all information available to us in a in a coherent and rational way.

$\longrightarrow$ Bayesian approach to probabilities offers a model for that.

But we also interact with others:

When I observe you acting in a certain way, or if I hear you expressing your probability estimate about a certain event, I might deduce from that some new information about the state of the world, which allows me to update the probabilities that I attribute to certain events.

And in turn: When you then see me acting in a certain way, or hear me expressing the probability that I attribute to a certain event (knowing that I have exploited all the information available to me), you might deduce from that some new information about the state of the world, which allows you to update the probabilities that you attribute to certain events.

$\longrightarrow$ We are engaged in a Bayesian dialogue.

Mathematicians working in economics, decision and game theory have studied such scenarios of indirect information exchange:

- DeGroot, 1974. "Reaching a consensus," *Journal of the American Statistical Association*.
- Dalkey, 1969. "The Delphi method: an experimental study of group opinion," *United States Air Force Project Rand*.
- Aumann, 1976. "Agreeing to disagree," *The Annals of Statistics* 4.
- Geanakoplos and Polemarchakis, 1982. "We can't disagree forever," *Journal of Economic Theory*.
- Bacharach, 1979. "Normal bayesian dialogues." *Journal of the American Statistical Association*.
- Fagin, Halpern, Moses, Vardi. 1995. *Reasoning about Knowledge.* MIT Press. (First published as *Reasoning about Knowledge*, *Mimeo, IBM*, San Jose, California, 1988.)

Forerunners:

- LACAN, 1945. "Le temps logique et l'assertion de certitude anticipée : un nouveau sophisme," *Les Cahiers d'Art, 1940–1944*; reprinted in: *Ecrits* (1966). Le Seuil: Paris.
- LITTLEWOOD, 1953. *A Mathematician's Miscellany*, Methuen & Co, London.

In markets:

- MILGROM and STOKEY. 1982 "Information, trade, and common knowledge," *Journal of Economic Theory*.
- SEBENIUS and GEANAKOPLOS. 1983. "Don't bet on it : contingent agreements with asymmetric information," *Journal of the American Statistical Association*.

More recent work:

- POLEMARCHAKIS, 2016. "Rational dialogs," Working Paper.

Part I

The formal framework (and Aumann's theorem)

The formal framework (Aumann 1976)

Let $(\Omega, \mathscr{B}, p)$ be a probability space:

- Ω the set of possible states of the world,
- $\mathscr{B}$ a σ-algebra on Ω, and
- p the prior probability distribution defined on $(\Omega, \mathscr{B})$.

Furthermore: Two individuals, 1 and 2, who impute the same prior probability p to the events in $\mathscr{B}$ but who have access to private information, given by a finite partition $\mathscr{P}_i$ of Ω, that is, a finite set

$$\mathscr{P}_i = \{P_{i1}, P_{i2}, \ldots, P_{ik}, \ldots, P_{iK_i}\}$$

of nonempty subsets of Ω, the *classes* of the partition, such that:

(a) each pair $(P_{ik}, P_{ik'})$, $k \neq k'$, is disjoint and

(b) $\bigcup_k P_{ik} = \Omega$.

The partition $\mathscr{P}_i$ models individual i's information: when $\omega \in \Omega$ is the true state, the individual characterized by $\mathscr{P}_i$ will learn that one of the states that belong to the class of the partition $\mathscr{P}_i$ to which belongs ω, which shall be denoted by in $P_i(\omega)$, has materialized.

In order to guarantee that the classes P_{ik} of the partition $\mathscr{P}_i$ are measured by p, we suppose, of course, that they belong to $\mathscr{B}$.

Example

Let $\Omega = \{a, b, c, d, e, f, g, h, i, j, k\}$ and

$$\mathscr{P}_i = \{\{a, b, g, h\}, \{c, d, i, j\}, \{e, f, k\}\}.$$

Assume $\omega^\star = c$, the true state of the world. Then individual i, modeled by the partition above, will only receive the information that the true state of the world is in $\{c, d, i, j\}$, that is, that one of the states in $\{c, d, i, j\}$ has materialized (but not which one exactly). In our notation: $P_i(c) = \{c, d, i, j\}$.

Important: We, as the theorist who builds the model, know that the true state is $\omega^\star = c$, but the individual in the model does not know it. He, or she, only knows that it is one of the states in $\{c, d, i, j\}$.

With this interpretation, if ω is the true state and $P_i(\omega) \subset A$, that is, $P_i(\omega)$ *implies* A, then individual i (at state ω) "knows" that event A has happened.

In the example above:

$$\mathscr{P}_i = \{\{a, b, g, h\}, \{c, d, i, j\}, \{e, f, k\}\}.$$

When $\omega^\star = c$ is the state that has materialized, individual i will *know* that the state that has materialized is in $\{c, d, i, j\}$, that is, that the event $\{c, d, i, j\}$ has occurred.

As a consequence, i will also know that any event that is a superset of $\{c, d, i, j\}$ has occurred. For example, i will also know that the event $\{a, c, d, i, j, e\}$ has occurred. And certainly, i will also know that any event disjoint of $\{c, d, i, j\}$ did not occur. For example, i will also know that the event $\{a, e\}$ *did not* occur.

A crucial assumption:

Following Aumann (1976), we assume that the prior p defined on $(\Omega, \mathscr{B})$ as well as the information partitions of the two individuals, $\mathscr{P}_i$, $i \in I = \{1, 2\}$, are *common knowledge* between the two individuals.

According to David Lewis (1969), an event is *common knowledge* between two individuals if not only both know it but also both know that the other knows it and that both know that the other knows that they both know it, ad infinitum (Lewis 1969).

A probabilistic model of "beliefs"

More generally, if individual i is Bayesian rational, then for any event A that belongs to the σ-algebra defined on Ω, after realization of the true state of the world, i can calculate the posterior probability of A given the information provided by the partition $\mathscr{P}_i$, that is, the *conditional probability of A given that the true state belongs to $P_i(\omega)$*:

$$q_i = p(A \mid P_i(\omega)) = \frac{p(A \cap P_i(\omega))}{p(P_i(\omega))}.$$

Remember: $P_i(\omega)$ is the information class (or "cell") of i's partition in which lies the state ω.

Example Let $\Omega = \{a, b, c, d, e, f, g, h, i, j, k, l, m\}$, endowed with uniform prior, that is, $p(\omega) = 1/13$ for all possible states of the world, and

$$\begin{aligned} \mathscr{P}_1 &= \{\{a, b, c, d, e, f\}, \{g, h, i, j, k\}, \{l\}, \{m\}\}, \\ \mathscr{P}_2 &= \{\{a, b, g, h\}, \{c, d, i, j\}, \{e, f, k\}, \{l, m\}\}. \end{aligned}$$

Let $A = \{a, b, i, j, k\}$ be the event of interest; and $\omega^\star = a$. Then:

$$\begin{aligned} q_1 = \frac{P(A \mid P_1(a))}{P_1(a)} &= \frac{p(\{a, b, i, j, k\} \cap \{a, b, c, d, e, f\})}{p(\{a, b, c, d, e, f\})} \\ &= \frac{p(\{a, b\})}{p(\{a, b, c, d, e, f\})} = \frac{1}{3} \\ q_2 = \frac{P(A \mid P_2(a))}{P_1(a)} &= \frac{p(\{a, b, i, j, k\} \cap \{a, b, g, h\})}{p(\{a, b, g, h\})} \\ &= \frac{p(\{a, b\})}{p(\{a, b, g, h\})} = \frac{1}{2} \end{aligned}$$

Terminology:

In game theory, decision theory, and economics, the probability attributed to an event is also called a *belief*.

In this terminology, $p(A)$ is the *prior belief of A*, which by assumption is common knowledge between the two individuals, and $p(A \mid P_i(\omega))$ the *posterior belief* that i attributes to A given the information received through his or her partition.

Remember: According to David Lewis (1969), an event is *common knowledge* between two individuals if not only both know it but also both know that the other knows it and that both know that the other knows that they both know it, ad infinitum (Lewis 1969).

To capture this notion within a set-theoretic framework that relies on the notion of a state of the world, it turns out to be useful—and having established this is one of Aumann's achievements—to consider the *meet* of the two partitions.

Definition 1 Let $\mathscr{P}_1$ and $\mathscr{P}_2$ be two partitions of Ω. The *meet* of $\mathscr{P}_1$ and $\mathscr{P}_2$, denoted by $\hat{\mathscr{P}} = \mathscr{P}_1 \wedge \mathscr{P}_2$, is the *finest common coarsening* of $\mathscr{P}_1$ and $\mathscr{P}_2$, that is, the finest partition of Ω such that, for each $\omega \in \Omega$,

$$P_i(\omega) \subset \hat{P}(\omega), \quad \forall i \in I = \{1, 2\},$$

where $\hat{P}(\omega) = P_1 \wedge P_2(\omega)$ is the class of the meet to which belongs ω.

Example

$$\begin{aligned} \mathscr{P}_1 &= \{\{a, b, c, d, e, f\}, \{g, h, i, j, k\}, \{l\}, \{m\}\}, \\ \mathscr{P}_2 &= \{\{a, b, g, h\}, \{c, d, i, j\}, \{e, f, k\}, \{l, m\}\}. \end{aligned}$$

$$\hat{\mathscr{P}} = \mathscr{P}_1 \wedge \mathscr{P}_2 = \{\{a, b, c, d, e, f, g, h, i, j, k\}, \{l, m\}\}$$

The meet of the two information partitions, casually speaking, represents what is *common knowledge* between the two individuals. The following definition makes this more precise.

Lemma (Aumann 1976) An event $A \subset \Omega$, at state ω, is *common knowledge* between individuals 1 and 2 in the sense of the recursive definition (Lewis 1969) if and only if $\hat{P}(\omega) \subset A$, that is, if the information class of the meet of the two partitions to which belongs the state ω is contained in A.

Example

$$
\begin{aligned}
\mathscr{P}_1 &= \{\{a, b, c, d, e, f\}, \{g, h, i, j, k\}, \{l\}, \{m\}\}, \\
\mathscr{P}_2 &= \{\{a, b, g, h\}, \{c, d, i, j\}, \{e, f, k\}, \{l, m\}\}.
\end{aligned}
$$

Remark. Of course, if $\mathscr{P}_1 \wedge \mathscr{P}_2$ is a coarsening of the two individual partition, then each of the individual partitions is a refinement of $\mathscr{P}_1 \wedge \mathscr{P}_2$: That is, if P is a class of the meet $\mathscr{P}_1 \wedge \mathscr{P}_2$, then, for each i, the union of th classes P_{ik} of the partition $\mathscr{P}_i$ contained in P is P,

$$\bigcup_{P_{ik} \subset P} P_{ik} = P,$$

and hence $\mathscr{P}_i$ induces a partition of P.

This is easy to verify in the example from above:

$$\begin{aligned} \mathscr{P}_1 &= \{\{a, b, c, d, e, f\}, \{g, h, i, j, k\}, \{l\}, \{m\}\}, \\ \mathscr{P}_2 &= \{\{a, b, g, h\}, \{c, d, i, j\}, \{e, f, k\}, \{l, m\}\}. \end{aligned}$$

$$\hat{\mathscr{P}} = \mathscr{P}_1 \wedge \mathscr{P}_2 = \{\{a, b, c, d, e, f, g, h, i, j, k\}, \{l, m\}\}$$

Example

$\Omega = \{a, b, c, d, e, f, g, h, i, j, k, l, m\}$, with uniform prior,

$$\begin{aligned} \mathscr{P}_1 &= \{\overbrace{\{a,b,c,d,e,f\}}^{1/3}, \overbrace{\{g,h,i,j,k\}}^{3/5}, \overbrace{\{l\}}^{0}, \overbrace{\{m\}}^{0}\}, \\ \mathscr{P}_2 &= \{\underbrace{\{a,b,g,h\}}_{1/2}, \underbrace{\{c,d,i,j\}}_{1/2}, \underbrace{\{e,f,k\}}_{1/3}, \underbrace{\{l,m\}}_{0}\}, \end{aligned}$$

Let $A = \{a, b, i, j, k\}$. For individual 1: attributing to A a posterior of $1/3$ corresponds to the event $\{a, b, c, d, e, f,\}$; attributing to A a posterior of 0 corresponds to the event $\{l, m\}$; attributing to A a nonzero posterior corresponds to the event $\{a, b, c, d, e, f, g, h, i, j, k\}$

For individual 2, attributing to A a posterior of $1/2$ corresponds to the event $\{a, b, c, d, g, h, i, j\}$, etc.

Common knowledge of posteriors

$$
\begin{aligned}
\mathscr{P}_1 &= \{\overbrace{\{a,b,c,d,e,f\}}^{1/3}, \overbrace{\{g,h,i,j,k\}}^{3/5}, \overbrace{\{l\}}^{0}, \overbrace{\{m\}}^{0}\}, \\
\mathscr{P}_2 &= \{\underbrace{\{a,b,g,h\}}_{1/2}, \underbrace{\{c,d,i,j\}}_{1/2}, \underbrace{\{e,f,k\}}_{1/3}, \underbrace{\{l, m\}}_{0}\},
\end{aligned}
$$

Suppose that $\omega^\star = m$ the true state of the world. Then, individual 1 will attribute to A a posterior of 0. This fact will be common knowledge between the two individuals, even though individual 2 does not know whether 1 has received the information that the true states belongs to $\{l\}$ or to $\{m\}$. This will be so, because for any of these two cases, individual 1 will always have calculated a posterior of 0.

At the same time, individual 2 will attribute to A a posterior of 0, and this will also be common knowledge.

Aumann's (1976) "agreement" result

Robert Aumann, (1976) "Agreeing to disagree," *The Annals of Statistics* 4 (6): 1236-1239.

- In economics, Aumann's paper has stimulated a rich literature.
- Derives its importance also for the formal framework that it proposes for modeling knowledge and common knowledge.
- What is this result?

Proposition (Aumann 1976)

Let $(\Omega, \mathscr{B}, p)$ a probability space, $\mathscr{P}_1$ and $\mathscr{P}_2$ two finite partitions of Ω, measurable with respect to $\mathscr{B}$, that represent the information accessible to individual 1 respectively 2, all of this being common knowledge between the two individuals. Let furthermore $A \in \mathscr{B}$ be an event. If at state ω (in virtue of the common knowledge of the prior probability and the information partitions) the posteriors q_1 and q_2 that the individuals attribute to A are common knowledge, then they have to be equal: that is, $q_1 = q_2$.

The proof

Can be understood in three steps. Step 1 (conceptually the most important) consists in establishing that common knowledge of q_i implies that for any information class of $\mathscr{P}_i$ that is a subset of the information class of the meet to which belongs the true state, $P_i(\omega)$, the conditional probability of A has to be equal to q_i:

$$q_i = \frac{p(A \cap P_i(\omega))}{p(P_i(\omega))} = \frac{p(A \cap P_{ik})}{p(P_{ik})}, \quad \forall\, P_{ik} \subset \hat{P}(\omega). \tag{1}$$

Otherwise there would be some level of knowledge at which q_i would not be known, and therefore cannot be common knowledge.

Illustration:

$$\begin{aligned} \mathscr{P}_1 &= \{ \overbrace{\{a, b\}}^{p(\{b,c\}|\{a,b\})=\frac{1}{2}}, \overbrace{\{c, d\}}^{p(\{b,c\}|\{c,d\})=\frac{1}{2}}, \{e\}, \{f\}\} \\ \mathscr{P}_2 &= \{ \{a, c\}, \{b, d\}, \{e, f\}\}, \end{aligned}$$

where $A = \{e, f\}$, and a the true state; $\hat{P}(a) = \{a, b, c, d\}$

Step 2: From (1) and the fact that the classes of i's partition that are subsets $\hat{P}(\omega)$ induce a partition of $\hat{P}(\omega)$, one obtains that:

$$q_i = \frac{p(A \cap \hat{P}(\omega))}{p(\hat{P}(\omega))}. \tag{2}$$

To see why (2) holds, note that (1) can be written as

$$p(A \cap P_{ik}) \;=\; q_i \, p(P_{ik}), \qquad \forall \, P_{ik} \subset \hat{P}(\omega).$$

Summing over all $P_{ik} \subset \hat{P}(\omega)$ gives

$$\sum_{P_{ik} \subset \hat{P}(\omega)} p(A \cap P_{ik}) \;=\; q_i \sum_{P_{ik} \subset \hat{P}(\omega)} p(P_{ik}).$$

Since the P_{ik} are disjoint (because they are elements of a partition), and the union over all P_{ik}-s that are subsets of $\hat{P}(\omega)$ gives $\hat{P}(\omega)$, by the property of σ-additivity of the probability measure p we have:

$$p(A \cap \hat{P}(\omega)) \;=\; q_i \, p(\hat{P}(\omega)).$$

Rearranging terms gives equation (2).

Step 2 relies on the more general fact that if A_k is a sequence of disjoint subsets of Ω and $p(B \mid A_k) = q$ for all k, then $p(B \mid \cup A_k) = q$, which is a simple consequence of the Kolmogorov Axioms.[1]

Illustration:

$$\begin{aligned} \mathscr{P}_1 &= \{ \overbrace{\overbrace{\{\, a,\ b \,\}}^{p(A|\{a,b\})=\frac{1}{2}} \ ,\ \overbrace{\{\, c,\ d \,\}}^{p(A|\{c,d\})=\frac{1}{2}}}^{p(A|\{a,b,c,d\})=\frac{1}{2}} , \{e\}, \{f\}\} \\ \mathscr{P}_2 &= \{\{\, a,\ c \,\}, \{\, b,\ d \,\}, \{e, f\}\} \end{aligned}$$

---

[1] In its more general form, namely that if $\mathscr{H}$ is a sub-σ algebra of $\mathscr{G}$, then

$$\mathbf{E}[\mathbf{E}(X \mid \mathscr{G}) \mid \mathscr{H}] = \mathbf{E}[X \mid \mathscr{H}]$$

this property is sometimes referred to as the *Tower Property*; see, for instance Williams (1991). I would like to thank Mathias Beiglböck and Daniel Toneian for having pointed this out to me.

Step 3: Finally, from the fact that (2) has to hold for each of the two individuals, one obtains that:

$$q_1 = \frac{p(A \cap \hat{P}(\omega))}{p(\hat{P}(\omega))} = q_2. \tag{3}$$

which concludes the proof.

Illustration:

$$\begin{aligned}
\mathscr{P}_1 &= \{ \overbrace{\overbrace{\{a, b\}}^{p(A|\{a,b\})=\frac{1}{2}}, \overbrace{\{c, d\}}^{p(A|\{c,d\})=\frac{1}{2}}}^{p(A|\{a,b,c,d\})=\frac{1}{2}}, \{e\}, \{f\}\} \\
\mathscr{P}_2 &= \{ \underbrace{\underbrace{\{a, c\}}_{p(A|\{a,c\})=\frac{1}{2}}, \underbrace{\{b, d\}}_{p(A|\{b,d\})=\frac{1}{2}}}_{p(A|\{a,b,c,d\})=\frac{1}{2}}, \{e, f\}\}
\end{aligned}$$

The Aumann conditions

Putting (1)–(3) together, one has:

$$q_i = \frac{p(A \cap P_i(\omega))}{p(P_i(\omega))} = \frac{p(A \cap P_{ik})}{p(P_{ik})} = \frac{p(A \cap \hat{P}(\omega))}{p(\hat{P}(\omega))} \quad \forall P_{ik} \subset \hat{P}(\omega), \quad \forall i \in$$

That is, for each i, the posterior attributed to A, given $P(\omega)$, has to be equal to:

(1) the posterior probability of A given any of the classes P_{ik} of i's partition that are contained in the class of the meet to which belongs the true state of the world $\hat{P}(\omega)$, and

(2) the posterior probability of A given $\hat{P}(\omega)$, that is, the element of the meet to which belongs ω.

I refer to equation (4) as the *Aumann conditions*.

Example (in which the Aumann conditions hold)

$$\begin{aligned} \mathscr{P}_1 &= \{\{a,b\},\{c,d\},\{e\},\{f\}\}, \\ \mathscr{P}_2 &= \{\{a,c\},\{b,d\},\{e,f\}\}, \end{aligned}$$

$A = \{b, c\}$ the event of interest, and $\omega = a$ the true state of the world. Uniform prior, that is, $1/6$ for each possible state of the word. Then:

$$\begin{aligned} q_1 &= \frac{p(A \cap P_1(a))}{p(P_1(a))} = \frac{p(\{b,c\} \cap \{a,b\})}{p(\{a,b\})} = \frac{p(\{b\})}{p(\{a,b\})} = \frac{1}{2} \\ q_2 &= \frac{p(A \cap P_2(a))}{p(P_2(a))} = \frac{p(\{b,c\} \cap \{c,a\})}{p(\{c,a\})} = \frac{p(\{c\})}{p(\{c,a\})} = \frac{1}{2} \end{aligned}$$

The meet is $\hat{\mathscr{P}} = \{\{a, b, c, d\}, \{e, f\}\}$. Hence, $\hat{P}(a) = \{a, b, c, d\}$. Here, each i thinks it possible that the other has received any of the classes in the others partition that are included in $\hat{\mathscr{P}} = \{\{a, b, c, d\}$. However:

$$\begin{aligned}
\frac{p(\{b, c\} \cap \{c, d\})}{p(\{c, d\})} &= \frac{p(\{c\})}{p(\{c, d\})} = \frac{1}{2}, \\
\frac{p(\{b, c\} \cap \{d, b\})}{p(\{d, b\})} &= \frac{p(\{b\})}{p(\{d, b\})} = \frac{1}{2}.
\end{aligned}$$

And, as it should be according to the Aumann conditions:

$$p(\{b, c\} \mid \hat{P}(a)) = \frac{p(\{b, c\} \cap \{a, b, c, d\})}{p(\{a, b, c, d\})} = \frac{p(\{b, c\})}{p(\{a, b, c, d\})} = \frac{1}{2}.$$

Illustration:

$$\begin{aligned}
\mathscr{P}_1 &= \{ \overbrace{\overbrace{\{a,\ b\}}^{p(A|\{a,b\})=\frac{1}{2}},\ \overbrace{\{c,\ d\}}^{p(A|\{c,d\})=\frac{1}{2}}}^{p(A|\{a,b,c,d\})=\frac{1}{2}}, \{e\}, \{f\}\} \\
\mathscr{P}_2 &= \{ \underbrace{\underbrace{\{a,\ c\}}_{p(A|\{a,c\})=\frac{1}{2}},\ \underbrace{\{b,\ d\}}_{p(A|\{b,d\})=\frac{1}{2}}}_{p(A|\{a,b,c,d\})=\frac{1}{2}}, \{e, f\}\}
\end{aligned}$$

Direct communication

Imagine that after realization of the true state of the world the two individuals communicate to each other the information class of his or her partition of which they have learned that the true state of the world belongs to it. Such an exchange of information can be referred to as one of direct communication (see, for instance, Geanakoplos and Polemarchakis 1982).

What the individuals know after such an exchange is given by the intersection of the two respective classes of their information partitions. Over the entire range of Ω, the so defined set of subsets of Ω is given by the coarsest common refinement of the two partitions: their so-called *join*.

Definition 2 Let $\mathscr{P}_1$ and $\mathscr{P}_2$ two partitions of Ω. The *join* of $\mathscr{P}_1$ and $\mathscr{P}_2$, denoted by $\check{\mathscr{P}} = \mathscr{P}_1 \vee \mathscr{P}_2$, is the *coarsest common refinement* of $\mathscr{P}_1$ and $\mathscr{P}_2$, that is, the coarsest partition of Ω such that, for each $\omega \in \Omega$,

$$\check{P}(\omega) \subset P_i(\omega), \quad \forall i \in I = \{1, 2\},$$

where $\check{P}(\omega) = P_1 \vee P_2(\omega)$ is the class of the join to which belongs ω.

The classes of the join are obtained by taking for each class of one partition its intersections with the classes of the other partition (see, for instance Barbut 1968).

In the Example from above:

$$\begin{aligned}
\mathscr{P}_1 &= \{\{a,b\},\{c,d\},\{e\},\{f\}\}, \\
\mathscr{P}_2 &= \{\{a,c\},\{b,d\},\{e,f\}\}, \\
\text{The join: } \check{\mathscr{P}} &= \{\{a\},\{b\},\{c\},\{d\},\{e\},\{f\}\} \\
\text{Remember, the meet: } \hat{\mathscr{P}} &= \{\{a,b,c,d\},\{e,f\}\}
\end{aligned}$$

A technical note: the matrix representation of two partitions

Any two finite partitions can be written in the form of a matrix such that

- the elements of the matrix are occupied by the elements of the join of the two partitions, with possibly some elements of the matrix empty but without any rows or columns completely empty, and
- the information classes of one individual correspond to the rows of the matrix and that of the other individual to the columns of the matrix (see, for instance, Barbut 1968).

In such a matrix, the classes of the meet of the two partitions appear as the unions of those elements of the join that have the same empty elements along rows as well as columns.

Example:

$$\begin{aligned}
\mathscr{P}_1 &= \{\{a,b\},\{c,d\},\{e\},\{f\}\},\\
\mathscr{P}_2 &= \{\{a,c\},\{b,d\},\{e,f\}\},\\
\text{the join: } \check{\mathscr{P}} &= \{\{a\},\{b\},\{c\},\{d\},\{e\},\{f\}\}\\
\text{the meet: } \hat{\mathscr{P}} &= \{\{a,b,c,d\},\{e,f\}\}
\end{aligned}$$

Practical for calculating the posteriors for a certain event A: Let $A = \{b, c\}$ and $\omega = a$ the true state of the world:

$\{a^\star\}$	$\{\mathbf{b}\}$		$\frac{1}{2}$
$\{\mathbf{c}\}$	$\{d\}$		$\frac{1}{2}$
		$\{e\}$	0
		$\{f\}$	0
$\frac{1}{2}$	$\frac{1}{2}$	0	

In the figure above, for each row, to the right of the vertical line (information class of individual 1), appears the conditional probability of A given that row; for each column, below the horizontal line (information class of individual 2), appears the conditional probability of A given the column.

Part II

Bayesian Dialogues

Geanakoplos and Polemarchakis's (1982) scenario of indirect communication

... Imagine that after having received their private information about the true state of the world (according to their information partition), the two individuals, turn in turn, communicate their posteriors back and forth, each round extracting the information that is contained in the announcement of the previous round.

This process is best understood as operating through a successive reduction of the set of possible states of the world:

- The process starts by discarding all states that are not in the information class of the meet to which belongs the true state of the world. Of course, because simply by having received the information through their partitions—thanks to the common knowledge of these partitions—it will be common knowledge between the two individuals that any state that is not in that class of the meet cannot be the true state of the world.
- Then, at each step t, with one of the individuals announcing the posterior probability that he or she attributes to the event of interest A at this step, it becomes common knowledge between the two individuals that a certain subset of Ω at step t cannot contain the true state of the world: namely the union of all those partition classes of the individual who has just announced his or her posterior that do not lead to that posterior. This subset is discarded from Ω at step t to give Ω at step $t + 1$.

More formally:

Let $\Omega_0 = \Omega$.
Step 1: $\Omega_1 = \hat{P}(\omega^\star)$, where $\omega^\star$ is the true state of the world.
Step t: $\Omega_t = \Omega_{t-1} \backslash \bar{\mathscr{P}}_{i(t-1),t-1}$, where

$$\bar{\mathscr{P}}_{i(t),t} = \bigcup_{i(t),k} P_{i(t),k}, \text{ such that } \frac{p(A \cap P_{i(t),k} \cap \Omega_t)}{p(P_{i(t),k} \cap \Omega_t)} \neq q_{i(t),t},$$

$$q_{i(t),t} = \frac{p(A \cap P_i(\omega) \cap \Omega_t)}{p(P_i(\omega) \cap \Omega_t)}$$

with $i(t)$ given by the sequence $1, 2, 1, 2, \ldots$ if individual 1 starts, and by $2, 1, 2, 1 \ldots$ if individual 2 starts.

The process ends—more precisely, will have reached an absorbing state—when a subset of Ω is reached such that the announcement of the posterior of any of the two individuals does not allow them to discard any more states.

This terminal subset of Ω will be one on which the "Aumann" conditions hold: the posteriors will be common knowledge—thanks to the common knowledge of the information partitions induced by the reduced set of states of the world at that step—and hence (as Aumann's result says) will be equal.

A dynamic foundation of Aumann's result

This process converges after a finite number of steps to a situation in which the posteriors are common knowledge and hence—by Aumann's result—identical (Geanakoplos et Polemarchakis 1982). In that sense, such a process can be interpreted as a dynamic foundation of Aumann's result.

If the Aumann conditions are satisfied on the original set Ω, the process stops immediately at step 1, or to say it more correctly, will have reached its absorbing state at step 1.

Important:

What stops a Bayesian dialogue are the "Aumann" conditions.

Depends on the order

A Bayesian dialog depends on the order in which the two individuals announce their posteriors (see, for instance, Polemarchakis 2016). Depending on whether it is individual 1 or individual 2 who starts the process, the process can end with *different* subsets of Ω.

Important:

On each of these two different terminal subsets of Ω, the "Aumann" conditions hold. The process, so to say, gets "stopped" by the Aumann conditions. But, on these two different terminal subsets of Ω, different posteriors attributed to A in common knowledge might arise.

An example (in which the order matters, derived from Polemarchakis, 2016)

Let $\Omega = \{a, b, c, d, e, f, g, h, i, j, k\}$, with uniform prior, $p(\omega) = 1/11$,

$$\begin{aligned} \mathscr{P}_1 &= \{\{a, b, c, d, e, f\}, \{g, h, i, j, k\}\}, \\ \mathscr{P}_2 &= \{\{a, b, g, h\}, \{c, d, i, j\}, \{e, f, k\}\}, \end{aligned}$$

$A = \{a, b, i, j, k\}$, the event of interest; and $\omega^\star = a$, the true state of the world.

In matrix representation:

$$\begin{array}{ccc|c} \{\mathbf{a}^\star, \mathbf{b}\} & \{c, d\} & \{e, f\} & \frac{1}{3} \\ \{g, h\} & \{\mathbf{i}, \mathbf{j}\} & \{\mathbf{k}\} & \frac{3}{5} \\ \hline \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & \end{array}$$

Note that: $\mathscr{P}_1 \wedge \mathscr{P}_2 = \{\Omega\}$, and
$\mathscr{P}_1 \vee \mathscr{P}_2 = \{\{a, b\}, \{c, d\}, \{a, b\}, \{e, f\}, \{a, b\}, \{g, h\}, \{i, j\}, \{k\}\}$.

In this example, the outcome of a Bayesian dialogue depends on the order in which the two individuals report their posteriors.

If individual 1 starts:

Step 1: $\Omega(1) = \{a, b, c, d, e, f, g, h, i, j, k\}$,
$\mathscr{P}_{1,\Omega(1)} = \{\{a, b, c, d, e, f\}, \{g, h, i, j, k\}\}$

$$q_1 = \frac{p(\{a, b, i, j, k\} \cap \{a, b, c, d, e, f\})}{p(\{a, b, c, d, e, f\})} = \frac{p(\{a, b\})}{p(\{a, b, c, d, e, f\})} = \frac{1}{3}$$

If individual 1 announces $1/3$, then it will become common knowledge between the two individuals that the true state cannot belong to the set $\{g, h, i, j, k\}$, and therefore this set should be deleted from what remains in the *fund of common knowledge*. The matrix becomes:

$\{\mathbf{a}^\star, \mathbf{b}\}$	$\{c, d\}$	$\{e, f\}$	$\frac{1}{3}$
1	0	0	

Step 2: $\Omega(2) = \{a, b, c, d, e, f\}$, $\mathscr{P}_{2,\Omega(2)} = \{\{a, b\}, \{c, d\}, \{e, f\}\}$

$$q_2 = \frac{p(\{a, b\} \cap \{a, b\})}{p(\{a, b\})} = \frac{p(\{a, b\})}{p(\{a, b\})} = 1.$$

If individual 2 announces 1, then it will be common knowledge between the two individuals that the true state of the world cannot be in $\{c, d, e, f\}$, and hence this set can be deleted in common knowledge. The matrix becomes:

$\{\mathbf{a}^\star, \mathbf{b}\}$	1
1	

Step 3: $\Omega(3) = \{a, b\}$, $\mathscr{P}_{1,\Omega(3)} = \{\{a, b\}\}$. Individual 1 announces also "1," and the process has reached its absorbing state. Note that on the set of states that are still alive at step 3, $\Omega(3) = \{a, b\}$, the Aumann conditions are trivially satisfied because the information partitions of the two individuals induced by $\Omega(3) = \{a, b\}$ are identical: $\mathscr{P}_{1,\Omega(3)} = \{\{a, b\}\} = \mathscr{P}_{2,\Omega(3)}$.

In this example, the element of the join to which belongs the true state of the world is also $\{a, b\}$. Direct communication will therefore also lead to a posterior of 1 attributed to A.

But if individual 2 starts:

Step 1: $\Omega(1) = \{a, b, c, d, e, f, g, h, i, j, k\}$,
$\mathscr{P}_{2,\Omega(1)} = \{\{a, b, g, h\}, \{c, d, i, j\}, \{e, f, k\}\}$

$$q_1 = \frac{p(\{a, b, i, j, k\} \cap \{a, b, g, h\})}{p(\{a, b, g, h\})} = \frac{p(\{a, b\})}{p(\{a, b, g, h\})} = \frac{1}{2}$$

$\rightarrow \{e, f, k\}$ can be deleted in common knowledge. But then the matrix is:

$\{\mathbf{a}^\star, \mathbf{b}\}$	$\{c, d\}$	$\frac{1}{2}$
$\{g, h\}$	$\{\mathbf{i}, \mathbf{j}\}$	$\frac{1}{2}$
$\frac{1}{2}$	$\frac{1}{2}$	

And the process of deletion ends here, with each of them announcing $1/2$ from this moment on, forever.

If individual 1 starts:

Step 1:

$\{\mathbf{a}^\star, \mathbf{b}\}$	$\{c, d\}$	$\{e, f\}$	$\frac{1}{3}$
$\{g, h\}$	$\{\mathbf{i}, \mathbf{j}\}$	$\{\mathbf{k}\}$	$\frac{3}{5}$
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	

Step 2:

$\{\mathbf{a}^\star, \mathbf{b}\}$	$\{c, d\}$	$\{e, f\}$	$\frac{1}{3}$
1	0	0	

Step 3:

$\{\mathbf{a}^\star, \mathbf{b}\}$	1
1	

If individual 2 starts:

Step 1:

$\{\mathbf{a}^\star, \mathbf{b}\}$	$\{c, d\}$	$\{e, f\}$	$\frac{1}{3}$
$\{g, h\}$	$\{\mathbf{i}, \mathbf{j}\}$	$\{\mathbf{k}\}$	$\frac{3}{5}$
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	

Step 2:

$\{\mathbf{a}^\star, \mathbf{b}\}$	$\{c, d\}$	$\frac{1}{2}$
$\{g, h\}$	$\{\mathbf{i}, \mathbf{j}\}$	$\frac{1}{2}$
$\frac{1}{2}$	$\frac{1}{2}$	

Further properties of a Bayesian dialogue

The visible trace of a Bayesian dialogue is the sequence of announced posteriors.

It can be that at level "nothing happens," in the sense that each of the individuals repeats for a certain number of rounds the same posterior, while in the background, nevertheless, the two individuals—in common knowledge—successively discard possible states of the world, namely all those of which it has become common knowledge, up to that step, that they cannot be the true state of the world.

Any regularities in the sequence of announced posteriors stemming from a Bayesian dialogue?

Polemarchakis (2016) has recently addressed the following question: *Is there any pattern in the sequence of announced probabilities that stem from a Bayesian dialogue?*

Polemarchakis shows that there isn't: that for any sequence of numbers strictly between 0 and 1, $q_1, q_2, q_3, q_4, \ldots, q_N$, one can find a set Ω of possible states of the world and two partitions such that that sequence is the visible trace of a Bayesian, or as Polemarchakis says, a "rational dialogue."

Part III

Is it always rational fo speak Bayesian ?

Suppose that the two individuals appearing in the example above are two professional chess players who have been thrown into prison. The director of the prison calls on both of them and announces:

> *"Here is Ω, here are your information partitions, here is the prior p and the event A that we are interested in (all as defined above in the presentation of Example 2). One state of the world will materialize and each of you will receive information according to his or her information partition. Then, I will ask one of you, in front of the other: What is the probability that you attribute to the event A? After his answer, I will ask the other: What is the probability that you attribute to the event A? After his answer, I will turn to the first again and ask: Now, has the event A occurred or not? If his answer is correct, then both of you will get free. If it is not correct, both of you will sit for the rest of your lives."*

The director first calls on individual 2, in front of individual 1, and asks his posterior. Individual 2 truthfully says "1/2." The director then calls on individual 1 to step forward and to report her posterior. Individual 1 says ...

What do you think that she says?

The argument

Realize that after individual 2's announcement (see right-hand panel in Figure 2, step 2), what is left of Ω and the two information partitions induced by that set is common knowledge between the two individuals, which is to say that it is common knowledge between the two individuals that from this moment onward, if they were to state their posteriors truthfully, they would forever be stuck with the answers "$1/2$" (even if they had as many more rounds to go as they wanted). However, at a second thought, there is a way out of that situation.

Since the original information partitions as well as the announcement of individual 2, at the first step, are common knowledge between the two individuals, it is also common knowledge between the two individuals:

(1) that it is the very announcement of individual 2, at the first step, that has brought them into this situation in which the current updated belief of individual 1, whose turn it is now, has no informational value anymore,

(2) but also that before that—before individual 2's announcement of "$1/2$" and the ensuing reduction of the fund of common knowledge—the posterior that individual 1 had then did have some informational value, because it was either $1/3$ or $1/5$ (see right-hand panel in Figure 2, step 1).

Imagine that you are individual 2 (the one who the director has asked first and who will also be asked at the third step) and that in response to the director's question to individual 1, at the second step, you hear individual 1 say: "$1/3$." You know that this *cannot* possibly be the truthful Bayesian update of individual 1 after you have made your announcement at the previous step, because you know that individual 1's Bayesian update at this step is $1/2$, and in fact this is common knowledge between the two of you. But it is also common knowledge between the two of you that before your announcement at the first step, $1/3$ was a possible truthful Bayesian response of individual 1, which corresponds to the fact that individual 1 has received the information that the true state of the world belongs to the set $\{a, b, c, d, e, f\}$ and hence cannot be in $\{g, h, i, j, k\}$ (see right-hand panel in Figure 2, step 1).

Knowing that individual 2 is highly rational (and does not say "1/3" because she made a mistake in determining her Bayesian posterior after your announcement), you will probably infer that this—that at step 1, her Bayesian posterior was $1/3$—is precisely what individual 1 wants to tell you by her announcement of $1/3$. What you do, so to say, is to look for some kind of repair, some way to reconcile what you observe (which is commonly known *not* to be a truthful Bayesian response) with strategically rational behavior, and you understand that the states in $\{g, h, i, j, k\}$ can be discarded from the set of possible states of the world, which leaves you with $\{a, b, c, d\}$ as the fund of common knowledge at this step of your "conversation." Combining that with your own information, that the true state belongs to $\{a, b, g, h\}$, you understand that the true state belongs to $\{a, b\}$ and that the event $A = \{a, b, i, j, k\}$ has hence surely occurred. You announce "1," and both of you get free.

Now that was a thought experiment of individual 2 at the last step. If you are individual 1, you understand that this is the way that individual 2 will reason. Anticipating this, you as individual 1, at the second step, announce "$1/3$."

Both philosophers of language and game theorists might cry out and say: “Sure! But we have long had a name for this!”:

“This is ...

... “a conversational implicature”!
... “forward induction”!

A linguistic interpretation: a "Bayesian implicature"

In the story above, the profitable deviation from truthfully stating one's Bayesian updated belief thrives on the fact that *by doing the deviation*, it will become *common knowledge* that the announced probability cannot possibly be the speaker's Bayesian updated belief that this step; in other words, that she has deviated from the rule of truthfully stating her Bayesian updated belief at that step.

Philosophers of language and linguists might recognize in this movement a *conversational implicature* (Grice, 1975): the phenomenon that the meaning of a speech act, here the announced probability, will be implied by a deviation from some predefined convention how to talk under normal conditions—what Grice calls the “flouting” of a conversational maxim.

Under the name of the "Cooperative Principle," Grice isolates four main conversational maxims, "supermaxims," as he says:

- The maxim of Quantity: "1. Make your contribution as informative as required (for the current purpose of the exchange). 2. Do not make your contribution more informative than is required."
- The maxim of Quality: "Try to make your contribution one that is true."
- The maxim of Relation: "Be relevant."
- The maxim of Manner: "Be perspicuous."

Under the category of quality, Grice places two submaxims: "1. Do not say what you belief to be false." "2. Do not say that for which you do lack adequate evidence." Under the category of manner, Grice places: "1. Avoid obscurity of expression." "2. Avoid ambiguity." "3. Be brief (avoid unnecessary prolixity)." "4. Be orderly."

In the example above, the maxim flouted can be said to be that of quality, which here takes the specific form that one ought to truthfully announce one's Bayesian updated belief at the current state of a conversation. In other words, the maxim to truthfully report one's Bayesian updated belief at the current state of a conversation can be seen a submaxim of the supermaxim of quality. The implicature comes from it being common knowledge that a deviation from that maxim has occurred because the probability stated cannot possibly be the speaker's truthful Bayesian updated belief at that step. To "flout" a maxim, as Grice explains, is to "blatantly fail to fulfill it" (49). But what can be a more blatantly committed offense than one that is committed *in the face of common knowledge*? We have here indeed a mathematically precise manifestation of a communicative implicature. I propose to call such an implicature that thrives on it being common knowledge that an expressed belief cannot possibly be the truthful Bayesian updated belief of the speaker (at the current state of the conversation), a *Bayesian implicature*.

In addition to that one could bring to bear that the implicature observed in the example above is triggered by a *clash* of maxims (Grice 1975, 49). In a situation in which the Bayesian updated beliefs of the two individuals are already common knowledge (given the fund of common knowledge Ω_t at the current step t), reporting one's Bayesian updated belief would amount to making a perfectly *irrelevant* speech act. In other words, there is a clash between the maxim of *quality* and that of *relation*. In the example above, individual 1, when at step 2 she announces her original posterior, can be said to sacrifice *quality* in order to save *relevance*.

Additional material

Example (Geanakoplos et Polemarchakis 1982; attributed to Aumann)

In general parametric form, for any n. Here we see the case $n = 3$.

In matrix form:

$\{\mathbf{a}^\star, b, c\}$			$\frac{1}{3}$
$\{d\}$	$\{\mathbf{e}, f\}$		$\frac{1}{3}$
	$\{g, h\}$	$\{\mathbf{i}\}$	$\frac{1}{3}$
$\frac{1}{4}$	$\frac{1}{4}$	1	

References

AUMANN, Robert J. 1976. "Agreeing to disagree," *The Annals of Statistics* 4, 1236–1239.

BACHARACH, Michael. 1979. "Normal bayesian dialogues." *Journal of the American Statistical Association* 74, 837–846.

BARBUT M. 1968. "Partitions d'un ensemble fini: leur treillis (cosimplexe) et leur représentation géométrique." *Mathématiques et Sciences Humaines* 22, 5–22.

BILLOT, Antoine. 2007. "Le raisonnement stratégique (Ce qu'Aumann doit à Lacan)." in: Bernard Walliser (Ed.) *Economie et Cognition*, Editions Ophrys, 115–146.

DEGROOT, Morris H. 1974. "Reaching a consensus," *Journal of the American Statistical Association* 69, 118–121.

Dalkey, Norman C. 1969. "The Delphi method: an experimental study of group opinion," *United States Air Force Project Rand* RM-5888-PR: 1–79.

Fagin, Ronald, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. 1995. *Reasoning about Knowledge.* Cambridge, MA: MIT Press. (First published as *Reasoning about Knowledge*, *Mimeo, IBM*, San Jose, California, 1988.)

Geanakoplos, John. 1992. "Common knowledge," *Journal of Economic Perspectives* 6 (4), 53–82.

Geanakoplos, John and Herakles M. Polemarchakis. 1982. "We can't disagree forever," *Journal of Economic Theory* 28, 192–200.

Lacan, Jaques 1945. "Le temps logique et l'assertion de certitude anticipée : un nouveau sophisme," *Les Cahiers d'Art, 1940–1944*, 32–42; reprinted in: *Ecrits* (1966). Le Seuil: Paris.

Lewis, David K. 1969. *Convention: A Philosophical Study.* Cambridge University Press, Cambridge, MA.

Littlewood, J. E. 1953. *A Mathematician's Miscellany*, Methuen & Co, London.

Ménager, Lucie. 2006. "Connaissance commune et consensus," *Revue d'Economie Industrielle* 114: 41–66.

Milgrom, Paul and Nancy Stokey. 1982 "Information, trade, and common knowledge," *Journal of Economic Theory* 26 (1): 17–27.

Nielsen, Lars Tyge. 1984. "Common knowledge, communication, and convergence of beliefs," *Mathematical Social Sciences* 8, 1–14.

Polemarchakis, Herakles. 2016. "Rational dialogs." Working Paper (February 2016)

Sebenius, James K. and John Geanakoplos. 1983. "Don't bet on it : contingent agreements with asymmetric information," *Journal of the American Statistical Association* 78 (382): 424–426.

WILLIAMS, David. 1991. *Probability with Martingales*. Cambridge, UK: Cambridge University Press.