Modelling Knowledge

A model of Knowledge

Ingredients in a model of knowledge:

- states Ω,
- information function describing what an agent knows,
- a notion of knowledge.

States

- A state describes contingencies that are relevant for a particular decision.

- Here, a state is a complete description of the world, including an agent's information, beliefs, and behavior.

Information

We define:

Definition. An information function for Ω is a function h that associates with each state $\omega \in \Omega$ a non-empty subset $h(\omega)$ of Ω.

We interpret $h(\omega)$ as the set of states that an agent considers possible at ω.

Definition. An information function is *partitional* if there is some partition of Ω such that for any $\omega \in \Omega$, $h(\omega)$ is the element of the partition that contains ω.

An information function is partitional if and only if it satisfies the following two properties:

P1 $\omega \in h(\omega)$ for every $\omega \in \Omega$.

P2 If $\omega' \in h(\omega)$, then $h(\omega') = h(\omega)$.

Given some state ω, Property P1 says that the agent is not convinced that the state is not ω.

Property P2 says that:

- if ω' is also deemed possible,
- then the set of states that would be deemed possible were the state actually ω' must be the same as those currently deemed possible at ω.

Example 1. Suppose: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and that the agent's partition is $\{\{\omega_1, \omega_2\}, \{\omega_3\}\{\omega_4\}\}$. Then $h(\omega_3) = \{\omega_3\}$, while $h(\omega_1) = \{\omega_1, \omega_2\}$.

Example 2. Suppose $\Omega = \{\omega_1, \omega_2\}$, $h(\omega_1) = \{\omega_1\}$ but $h(\omega_2) = \{\omega_1, \omega_2\}$.

Here h is not partitional.

Knowledge

We will say that an agent knows E if E obtains at all the states that the agent believes are possible.

We refer to a set of states $E \subset \Omega$ as an event:

- if $h(\omega) \subset E$, then in state ω, the agent views $\neg E$ as impossible.

- Hence we say that the agent knows E at ω.

We define the agent's knowledge function K by:

$$K(E) = \{\omega \in \Omega : h(\omega) \subset E\}.$$

$K(E)$ is the set of states at which the agent knows E.

Example 1. cont. In our first example, agent's partition is $\{\{\omega_1, \omega_2\}, \{\omega_3\}\{\omega_4\}\}$.

Suppose $E = \{\omega_3\}$. Then $K(E) = \{\omega_3\}$.

Similarly, $K(\{\omega_3, \omega_4\}) = \{\omega_3, \omega_4\}$ and $K(\{\omega_1, \omega_3\}) = \{\omega_3\}$.

Example 2. cont. In our second example, $h(\omega_1) = \{\omega_1\}$ but $h(\omega_2) = \{\omega_1, \omega_2\}$.

So $K(\{\omega_1\}) = \{\omega_1\}$, $K(\{\omega_2\}) = \emptyset$ and $K(\{\omega_1, \omega_2\}) = \{\omega_1, \omega_2\}$.

Axioms of knowledge

Now, notice that for every state $\omega \in \Omega$, we have $h(\omega) \subset \Omega$. Therefore it follows that:

K1 (Axiom of Awareness) $K(\Omega) = \Omega$.

Regardless of the actual state, the agent knows that he is in some state.

K2 $K(E) \cap K(F) = K(E \cap F)$.

Property K2 says that if the agent knows E and knows F, then he knows $E \cap F$.

An implication of this property is that

$$E \subset F \Rightarrow K(E) \subset K(F)$$

This means that if F occurs whenever E occurs, then knowing F means knowing E as well.

To see why this additional property holds, suppose $E \subset F$. Then $E = E \cap F$, so $K(E) = K(E \cap F)$.

Applying K2 implies that:

$$K(E) = K(E \cap F) = K(E) \cap K(F) \Rightarrow K(E) \subset K(F)$$

from which the result follows.

If the information function satisfies **P1**, the knowledge function also satisfies a third property:

K3 (Axiom of Knowledge) $K(E) \subset E$.

This says that if the agent *knows* E, then E *must have occurred*.

The agent cannot know something that is false.

Finally, if the information function is partitional (i.e. satisfies both **P1** and **P2**), the knowledge function satisfies two further properties:

K4 (Axiom of Transparency) $K(E) \subset K(K(E))$.

Property K4 says that if the agent knows E, then he knows that he knows E.

To see this, note that if h is partitional, then $K(E)$ is the union of all partition elements that are subsets of E.

Moreover, if F is any union of partition elements $K(F) = F$ (so actually $K(E) = K(K(E))$).

K5 (Axiom of Wisdom) $\Omega \backslash K(E) \subset K(\Omega \backslash K(E))$.

Property K5 states the opposite: if the agent does not know E, then he *knows that he does not know* E.

This means that the agent:

- understands and is aware of all possible states;
- and can reason based on states that might have occurred, not just those that actually do occur.

Common Knowledge

Suppose there are I agents with partitional information functions $h_1, \ldots, h_I$ and associated knowledge functions $K_1, \ldots., K_I$.

We say that an event $E \subset \Omega$ is *mutual knowledge* in state ω if it is known to all agents, i.e. if:

$$\omega \in K_1(E) \cap K_2(E) \cap \ldots \cap K_I(E) \equiv K^1(E)$$

An event E is common knowledge in state ω if it is known to everyone, everyone knows this, and so on.

Definition *An event* $E \subset \Omega$ *is common knowledge in state* ω *if:*

$$\omega \in K^1(E) \cap K^1K^1(E) \cap K^1K^1K^1(E) \cap \ldots$$

A definition of common knowledge also can be stated in terms of information functions.

Let us say that an event F is *self-evident* if for all $\omega \in F$ and $i = 1, \ldots, I$, we have $h_i(\omega) \subset F$.

Definition. *An event $E \subset \Omega$ is common knowledge in state $\omega \in \Omega$ if there is a self-evident event F for which $\omega \in F \subset E$.*

Example 3. Suppose $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and there are two individuals with information partitions:

$$H_1 = \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4\}\}$$

$$H_2 = \{\{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4\}\}$$

We can see that event $E = \{\omega_1, \omega_2\}$ is never common knowledge since E does not contain any self-evident event.

Moreover, not according to the first either since, $K_1(E) = \{\omega_1\}$ implies that $K_2K_1(E) = \emptyset$.

Note, however, that $K_1(E) = \{\omega_1\} \subset \{\omega_1, \omega_2\} = E$, so E is mutual knowledge at ω_1.

Is $F = \{\omega_1, \omega_2, \omega_3\}$ common knowledge?

Yes, since F is self-evident.

Moreover, since $K_1(F) = F$ and $K_2(F) = F$, it easy to check the first definition as well.

Lemma. *The following are equivalent:*

- $K_i(E) = E$ *for all* i,
- E *is self-evident,*
- E *is a union of members of the partition induced by* h_i *for all* i.

Proof. To see (i) and (ii) are equivalent, note that F is self-evident iff $F \subset K_i(F)$ for all i.

By **K4**, $K_i(F) \subset F$ always, so F is self-evident iff $K_i(F) = F$ for all i.

To see (ii) implies (iii), note that if E is self-evident, then $\omega \in E$ implies $h_i(\omega) \subset E$, so $E = \cup_{\omega \in E} h^i(\omega)$ for all i.

Finally (iii) implies (i) immediately. $\blacksquare$

Proposition *The two definitions of common knowledge are equivalent*.

Proof. Assume E is common knowledge at ω according to the first definition.

Then $E \supset K^1(E) \supset K^1K^1(E) \supset \ldots$ and ω is a member of each of these sets, which are thus non-empty.

Since Ω is finite, there is some set $F = K^1 \ldots K^1(E)$ for which $K_i(F) = F$ for all i.

So this set F, with $\omega \in F \subset E$ is self-evident and E is CK by the second definition.

Assume E is common knowledge at ω according to the second definition.

There is a self-evident event F with $\omega \in F \subset E$. Then $K_i(F) = F$ for all i.

So $K_1(F) = F$, and $K_1(F)$ is self-evident. Iterating this argument, $K_1 \ldots K_1(F) = F$ and each is self-evident.

Now $F \subset E$, so by **K2**, $F \subset K_1 \ldots K_1(E)$.

Since $\omega \in F$, E is CK at ω by the first definition.

Aumann (1976)'s Agreement Theorem

Could two individuals who share the same prior ever agree to disagree?

That is, if i and j share a common prior over states, could a state arise at which it was commonly known that:

- i assigned probability η_i to some event,
- j assigned probability η_j to that same event,
- and $\eta_i \neq \eta_j$.

Aumann characterized conditions under which this sort of disagreement is impossible.

Formally, let p be a probability measure on Ω, interpreted as the agents' prior belief.

For any state ω and event E, let $p(E|h_i(\omega))$ denote i's posterior belief, so $p(E|h_i(\omega))$ is obtained by Bayes' rule.

The event that “i assigns probability η_i to E” is:

$$\{\omega \in \Omega : p(E|h_i(\omega) = \eta_i\}.$$

Proposition *Suppose two agents have the same prior belief over a finite set of states* Ω*.*

- *If each agent's information function is partitional;*
- *and it is common knowledge in some state* $\omega \in \Omega$ *that agent 1 assigns probability* η_1 *to some event E and agent 2 assigns probability* η_2 *to E:*

then $\eta_1 = \eta_2$.

Proof. If the assumptions are satisfied, then there is some self-evident event F with $\omega \in F$ such that:

$$F \subset \left\{ \begin{array}{c} \{\omega_0 \in \Omega : p(E|h_1(\omega_0) = \eta_1\} \\ \cap\{\omega_0 \in \Omega : p(E|h_2(\omega_0)) \end{array} \right\}$$

Moreover, F is a union of members of i's information partition, say $\cup_k A_k^1$ for 1 and $\cup_k A_k^2$ for 2.

Since Ω is finite, let

$$F = \bigcup_k A_k^1 = \bigcup_k A_k^2.$$

Now, for any non-empty disjoint sets C, D with

$$p(E|C) = \eta_i \text{ and } p(E|D) = \eta_i,$$

we must have

$$p(E|C \cup D) = \eta_i.$$

For each k, $p(E|A_k^1) = \eta_1$, then:

$$p(E|F) = p(E|\bigcup_k A_k^1) = \eta_1$$

and similarly:

$$\eta_1 = p(E|F) = p(E|\bigcup_k A_k^2) = p(E|A_k^2) = \eta_2$$

thus implying $\eta_1 = \eta_2$.

The no trade theorem

Suppose there are two agents.

Let Ω be a set of states and X a set of consequences (trading outcomes).

Definition. *A contingent contract is a function mapping* Ω *into* X.

Let A be the space of contracts: $a : \Omega \to X$.

Each agent has a utility function $u_i : X \times \Omega \to R$.

Let $U_i(a) = u_i(a(\omega), \omega)$ denote i's utility from contract.

Note that $U_i(a)$ is a random variable that depends on the realization of ω.

Let $E[U_i(a)|H_i]$ denote i's expectation of $U_i(a)$ conditional on his information H_i.

A contingent contract b is *ex ante efficient* if there is no contract a such that, for both i, $E[U_i(a)] > E[U_i(b)]$.

Milgrom and Stokey's (1982) no trade theorem:

Proposition *If a contingent contract* b *is ex ante efficient, then it cannot be common knowledge between the agents that every agent prefers contract a to contract* b.

Proof. The claim is that there cannot be a state ω, which occurs with positive probability, at which the set:

$$E = \{\omega : E[U_i(a)|h_i(\omega)] > E[U_i(b)|h_i(\omega)] \text{ for all } i\}$$

if common knowledge.

Suppose to the contrary that there was such a state ω and hence a self-evident set F such that $\omega \in F \subset E$.

By the definition of a self-evident set, for all $\omega' \in F$ and all i, $h_i(\omega') \in F$.

So for all $\omega' \in F$ and all i:

$$E[U_i(a) - U_i(b)|h_i(\omega')] > 0.$$

Using the fact that i's information is partitional, we know that:

$$F = h_i(\omega_1) \cup h_i(\omega_2) \cup \ldots \cup h_i(\omega_n)$$

for some set of states $\omega_1, \ldots, \omega_n \in F$ and all i. (We can choose these states so that $h(\omega_k) \cap h(\omega_l) = \emptyset$).

It follows that for all i:

$$E[U_i(a) - U_i(b)|F] > 0.$$

That is: contract a strictly Pareto dominates contract b conditional on the event F.

But now we have a contradiction to the assumption that b is ex ante efficient, because it is possible to construct a better contract c by defining c to be equal to a for all states $\omega \in F$ and equal to b for all states $\omega \notin F$.

The electronic mail game

The issues that may arise with knowledge in strategic environments are illustrated by this game.

They play game G_b with probability $p < 1/2$ and G_a with probability $1 - p$.

In both games, players have two actions A, B.

	A	B
A	M, M	$1, -L$
B	$-L, 1$	$0, 0$

G_a (prob. $1-p$)

	A	B
A	$0, 0$	$1, -L$
B	$-L, 1$	M, M

G_b (prob $p<\frac{1}{2}$)

In both games it is mutually beneficial for the players to choose the same action.

But the action that is best depends on the game:

- in G_a the outcome $(A; A)$ is best;
- in game G_a the outcome $(B; B)$ is best.

In the payoffs $L > M > 1$.

Even if a player is sure that the game is G_b, it is risky for him to choose B unless he is sufficiently confident that his partner is going to choose B as well.

Player 1 know the true game.

Player b receives an informative signal $\{a, b\}$.

Assume first 2’s signal is uninformative.

Then there is a unique equilibrium in which both players play A.

The payoff is $(1-p)M$.

Assume now that the players can communicate.

1 sends a message to b

Now both players know the game

They play A in G_a and B in G_b

Consider now a variant.

Players can communicate, but imperfectly.

Specifically, the players are restricted to communicate via computers under the following protocol:

- If the game is G_b then player 1's computer automatically sends a message to player 2's computer;
- if the game is G_a then no message is sent.
- If a computer receives a message then it automatically sends a confirmation.
- This is so not only for the original message but also for the confirmation, the confirmation of the confirmation, and so on.

- There is a small probability $\varepsilon > 0$ that any given message does not arrive at its intended destination.
- If a message does not arrive then the communication stops.
- At the end of the communication phase each player's screen displays the number of messages that his machine has sent.

We need to specify a set of states and the players' information functions.

Define the set of states to be:

$$\Omega = \Big\{(Q_1; Q_2)\ :\ Q_1 = Q_2 \text{ or } Q_1 = Q_2 + 1\Big\}.$$

In the state $(q;q)$:

- player 1's computer sends q messages, all of which arrive at player 2's computer,
- and the qth message sent by player 2's computer goes astray.

In the state $(q+1; q)$:

- player 1's computer sends $q + 1$ messages,
- and all but the last arrive at player 2's computer.

Player 1's information function is defined by:

$$P_1(q,q) = \begin{cases} (q,q),(q,q-1) & \text{if } q > 0 \\ (0,0) & q = 0 \end{cases}$$

Player 2's information function is defined by:

$$P_1(q,q) = \{(q,q),(q+1,q)\} \text{ for all } q.$$

The game that is played is G_a if $(0,0)$, else it is G_b.

Player 1 knows the game in all states.

Player 2 knows the game in all states except $(0,0)$ and $(1,0)$.

In each of the states $(1,0)$ and $(1,1)$ player 1 knows that the game is G_b but does not know that player 2 knows it.

Similarly in each of the states $(1,1)$ and $(2,1)$ player 2 knows that the game is G_b but does not know whether player 1 knows that player 2 knows that the game is G_b.

And so on.

In any state $(q;q)$ or $(q+1;q)$ the larger the value of q the "closer" we are to common knowledge (or so it would seem).

We can study this as a Bayesian game.

The state set is:

$$\Omega = \left\{ (Q_1; Q_2) \ : \ Q_1 = Q_2 \text{ or } Q_1 = Q_2 + 1 \right\}.$$

The signal function i of each player i is defined by:

$$\tau_i(Q_1; Q_2) = Q_i.$$

Each player's belief on Ω is the same, derived from the technology and the assumption that prior p:

$$
\begin{aligned}
p_i(0,0) &= 1 - p \\
p_i(q+1,q) &= p\varepsilon(1-\varepsilon)^{2q} \\
p_i(q+1,q+1) &= p\varepsilon(1-\varepsilon)^{2q+1}
\end{aligned}
$$

The game $G((Q_1;Q_2))$ determines the payoffs.

Proposition. *The electronic mail game has a unique Nash equilibrium, in which both players always choose A.*

Proof. We will proceed by induction starting from the state 0,0.

In the state $(0,0)$ the action A is strictly dominant for player 1.

So that in any Nash equilibrium player 1 chooses A when receiving the signal 0.

If player 2 gets no message (i.e. his signal is 0) then he knows that:

- either player 1 did not send a message (an event with probability $1 - p$),
- or the message that player 1 sent did not arrive (an event with probability εp).

If player 2 chooses A then 2's expected payoff is at least:

$$\frac{(1-p)M}{1-p+\varepsilon p}$$

no matter what 1's choice is in $(1,0)$.

If player 2 chooses B then his payoff is at most:

$$\frac{-L(1-p)+p\varepsilon M}{1-p+\varepsilon p}$$

Therefore it is strictly optimal for player 2 to choose A when his signal is 0.

Assume now that we have shown that for all $(Q_1; Q_2)$ with $Q_1 + Q_2 < 2q$ players 1 and 2 both choose A in any equilibrium.

Consider player 1's decision when she sends q messages.

1 is uncertain whether $Q_2 = q$ or $Q_2 = q - 1$.

The probability that she assigns to $Q_2 = q - 1$ is:

$$z = \frac{\varepsilon}{\varepsilon + (1-\varepsilon)\varepsilon} > \frac{1}{2}$$

Thus she believes that it is more likely that her last message did not arrive, than that player 2 got the message.

If she chooses B then her expected payoff is at most:

$$-Lz + (1 - z)M$$

since under the induction assumption, she knows that if $Q_2 = q - 1$ then 2 chooses A.

If she chooses A then her payoff is at least 0.

Given that $L > M$ and $z > 1/2$, her best action is A.

By a similar argument, if players 1 and 2 both choose A in any equilibrium for all (Q_1, Q_2) with $Q_1 + Q_2 < 2q + 1$ then player 2 chooses A when his signal is q.

Hence each player chooses A in response to every possible signal. ■