Knowledge and Common Knowledge

We will represent player i's knowledge using a partition H_i of a "state space" Ω:

When the true state is ω, player i knows that is in the element of his partition that contains ω; the elements of the partition are the states i considers possible. Call this set $h_i(\omega)$.

Implicit: the state space Ω all relevant uncertainty: the player's information/uncertainty about the state of nature, his information about other's information etc. Note also that since by definition $\omega \in h_i(\omega)$, player i always thinks that the true state is possible.

<u>Assume</u>: Ω is finite, there is a common prior p on Ω, all states have positive probability. (drop zero-probability states.)

Assuming finiteness makes the math a lot easier, but later we will need to deal with larger state spaces if only to understand how restrictive the finiteness assumption is.

Definition: “Player i knows E at ω” if $h_i(\omega) \subseteq E$.

$K_i(E) \equiv \{\omega \mid h_i(\omega) \subseteq E\}$:
this is the set of states where i knows *E*.

This definition satisfies the following properties.
(proof is HW)

"Necessitation": $K_i(\Omega) = \Omega$:

Player i always knows the state space. As a consequence, player i knows all statements that are true for every point in the state space, i.e. all tautologies.

$K_i(E) = K_i K_i(E)$
(i knows *E* if and only if she knows that she knows it. Implicitly players know their own information structure.)

"Introspection": $-K_i(-K_i(E)) \subseteq K_i(E)$

If you don’t know that you don’t know E, you know E. So players can’t be unaware of any possibilities.

Now define the event “everyone knows E”:

$$K_{\mathrm{I}}(E) = \cap K_i(E) = \{\omega \,|\, \cup_i h_i(\omega) \subseteq E\}$$

The event “i know that j knows E” is

$$\{\omega \,|\, h_i(\omega) \subseteq K_j(\omega)\}$$

This language/interpretation of the set implicitly suppose that player i “knows” the partitions of player j.

More generally, in interpreting the model, we will assume that all players know the state space and the structure of the model.

If they don’t, we add in more states.

Suppose that $\Omega = \omega', \omega''$ and $H_1 = H_2 = \{(\omega'), (\omega'')\}$: both players have the discrete partition.

What if we want to model the situation where both 1 and two know the state, but 1 doesn’t know if 2 knows it? Add a copy of each state, so the state space is now $\overline{\Omega} = (\omega', d), (\omega'', d), (\omega', t), (\omega'', t)$ and let the partitions be

$$H_1 = \left\{ \left((w', d), (w', t) \right); \left((w'', d), (w'', t) \right) \right\}$$
$$H_2 = \left\{ \left((\omega', d) \right); \left((\omega'', d) \right); \left((w', t), (\omega'', t) \right) \right\}$$

The event “everyone knows that everyone knows E” is

$$K_I^2(E) = \{\omega \mid \cup_i h_i(\omega) \subseteq K_I(E)\}.$$

Define $K_I^n(E) = \{\omega \mid \cup_i h_i(\omega) \subseteq K_I^{n-1}(E)\}$

$K_I^n(E) \subseteq K_I^{n-1}(E)$because beliefs are never false) .

$$K_I^\infty(E) \equiv \cap_n K_I^n(E)$$

Since the state space is finite, this intersection process ends after finitely many rounds:
there is a finite n such that $K_I^\infty(E) = K_I^n(E)$

<u>Definition :</u> *E is common knowledge at* ω if $\omega \in K_I^\infty(E)$.

When E is common knowledge, every statement of the form i knows that j knows that k and m know that.. E is true" is true.

Aumann 76 provides an alternative characterization of common knowledge in terms of the meets of the players partitions

partition H'' is a *coarsening* of partition H' if for every $\omega,\ h'(\omega) \subseteq h''(\omega)$:
H'' has bigger - or coarser- elements than H'

H is a *common coarsening* of H' and H'' if it is a coarsening of H' and also a coarsening of H''.

The meet M of the partitions $H_1,...,H_n$ is the *finest common coarsening:*

M is a common coarsening of the H_i and there is no other common coarsening M' with $M'(\omega) \subseteq M(\omega)$ for all ω and strict inclusion for at least one.

“Reachability Lemma”

$\omega' \in M(\omega)$ if there is a chain of states w=
$\omega = \omega_0, \omega_1, ...\omega_m = \omega'$ such that for each ω_k there is a player *i(k)* s.t.
$h_{i(k)}(\omega_k) = h_{i(k)}(\omega_{k+1})$:

Easy to see that if ω' reachable from ω, then ω' must be in $M(\omega)$:

$M(\omega)$has to include ω_1 because player i(0) cant tell it apart from ω and the meet has to be no finer than i(0)'s partition; so the meet has to include ω_2 because player i(1) can't tell it from ω_1
and so on.
The minimality requirement is what implies that the meet is exactly these states and no others.

Theorem Auman 1976 let M= meet of individual players' partitions.

Event E is common knowledge at ω iff $M(\omega) \subseteq E$.

Intuition: If ω' is reachable from ω, there is someone who thinks that there is someone who thinks that … who thinks ω' is possible, so E can't be common knowledge if it excludes ω'.
Conversely, from the reachability lemma, if $M(\omega) \subseteq E$, then everyone knows that everyone knows that… E is true, so E is c.k.

Example: "Dirty faces"
3 players, each face either dirty (1) or clean (0)
state is 3-bit binary number (1's face, 2's, 3's)

Version 1: each player only knows state of other players faces; so
H1= {(000,100);(001,101);(010,110);(011,111)}
H2={000,010);(001,011); (100,110);(101,111)}
H3= {000,001);…

E*= {not 000}= {at least 1 dirty face)

When is E* ck?

Two ways to compute this.

1. Easy way: compute the meet
Claim: Every state is reachable from every state so $M(\omega)=\Omega$ for all ω.
Start at 000->001->011->010->110->100->101->111

So when is E* ck??

2: Second way: compute $K_I(E^*)$, $K_I^2(E^*)$, etc.
$K_I(E^*)$= ? $K_I^2(E^*)$=? $K_I^3(E^*)$

Version 2:

A sage enters the room, and if all faces are clean, says so. Everyone knows this is how he behaves so if he doesn’t say anything everyone knows at least 1 dirty face.

H1= {(000), (100),(001,101);(010,110);(011,111)}, etc.

Now if E* occurs everyone knows its true:
$E^* = K_I(E^*)$. (E* is a "public event.")
then $K_I^2(E^*) = K_I(K_I(E^*)) = K_I(E^*) = E^*$:
public events are common knowledge henever they occur.

Next: embed the dirty faces story in a (far fetched) game

3 periods. Each period all players simultaneously decide whether or not to blush. Can only blush once.
Actions revealed at end of each round
payoffs
(-100 blush clean; -1 never blush, dirty; 1 never blush clean; δ^t blush at t, dirty.)

with these payoffs, players want to blush iff they think its very likely their face is dirty.

Faces are i.i.d. d dirty/clean probability 1/2,

Consider info structure 1- no sage-

Claim : the only NE is " never blush regardless of what you see."

Sketch of Proof: (a) check that this is NE. Don't see own face, learn nothing from opponents s play, and given prior best response is not blush.
(b) fix a Nash equilibrium, and let t0 be first date where anyone blushes pos prob. Haven't learned anything, so 1/2-not blush is optimal.

Now info structure 2:

If sage says "all clean" then no on blushes.
Suppose sage says "at least 1 dirty."

- If state has exactly 1 dirty face: dirty player blushes in period 1.
- If no one blushes in period 1: all know that at least 2 dirty. So in period 2, a player who sees exactly 1 clean face knows his own face is dirty. So if exactly 2 dirty- those 2 blush in period 2.
- If no blush in period 1 or period 2: Everyone knows all 3 dirty: all blush in period 3.

Discussion: Consider the state where all faces dirty. In this state, the sage's info isn't directly payoff relevant given players information, but it still changes play. The reason is that it takes something that each player used to know but wasn't common knowledge and makes it c.k. But note that the sage's information *is* payoff relevant in states with exactly 1 dirty face- here it tells the player with the dirty face that his face is indeed dirty; which is what lets players infer something when the player doesn't blush. This is true in general: if a change in the information structure makes something c.k. at a state where it didn't use to be, there must be some state where some player's knowledge is different.

Theroem Aumann 1976: Suppose it is c.k. at ω that i's posterior probability of E is q_i and j's is q_j. Then $q_i = q_j$.

Informally, i and j can't "agree to disagee."
Of course if they get different signals they can have different beliefs; the idea is to use c.k. posteriors as a way to formalize "agreement".

Proof: The event that player *i's* posterior is q_i is $M(\omega) \subseteq$ player i's posterior is
$\{\omega \mid p(E \mid h_i(\omega)) = q_i\}$.
So it is common knowledge that player *i's* posterior is q_i when $M(\omega) \subseteq \{\omega \mid p(E \mid h_i(\omega)) = q_i\}$

Write the meet at ω as $M(\omega) = \bigcup_k h_i^k$, where each h_i^k is an element of H_i.
(One of these h_i^k will be $h_i(\omega)$; the other elements, if any, are in the meet because they are reachable from ω.)

Since q_i is constant on $M\left(\omega\right)$, $q_i = \frac{p(E \cap h_k^i)}{p(h_i^k)}$ for all k. Hence $p(E\bigcap h_i^k) = q_i\, p(h_i^k)$.

Now sum over k; since the h_i^k are disjoint, the sum of the probabilities is the probability of the union;

$$p(E \cap M(\omega)) = q_i\, p(M(\omega)).$$

The same reasoning for j shows

$$p(E \cap M(\omega)) = q_j\, p(M(\omega))$$

so $q_i = q_j$.

Amazingly slick. Where"d that rabbit come from?

Discussion

- Theorem relies on a common prior

- Interpretation as c.k. requires that the state encodes all of the players information: If i doesn't know j's partition, i doesn't know how to interpret the fact that j's posterior is different than his own.
- The theorem requires that the posteriors are ck, and not just that each player knows the posterior beliefs of the others:

$\Omega = \omega_1, \omega_2, \omega_3, \omega_4$, uniform prior p
$H_1 = \{(\omega_1, \omega_2), (\omega_3, \omega_4)\}$
$H_2 = \{(\omega_1, \omega_2, \omega_3), (\omega_4)\}$
$E = (\omega_1, \omega_4)$

$q_1(E \mid \omega) = 1/2 \forall \omega$, so 2 knows $q_1(E)$.

$q_2(E \mid \omega_1) = q_2(E \mid \omega_2) = 1/3.$
So at ω_1 and ω_2, 1 knows $q_2(E)$.
So each player knows other player's posterior, but the posteriors aren't equal. Of course posteriors not c.k.

The theorem doesn’t explain how posteriors get to be ck.

Geanakoplos-Polermarchakis "We cannot disagree forever" : Players take turns truthfully reporting their posteriors. (no consideration of incentives.)

<u>Result</u>: if state space is finite, process converges to stationary point in finite time- at this point posteriors are ck.

At the stationary point, players need not know as much as if they pooled their information.
Counterexample: Take an event where each player’s posterior is always ½., but pooling signals is fully revealing.

The “agreeing to disagree” theorem is closely related to no-speculation theorems

Suppose we are both risk averse, and bet on a coin flip: I win a dollar from you if H, you win a dollar from me if T. In a sense we are agreeing to disagree: you must have a posterior of H <1/2 and mine must be 1/2.

--

How does set of NE change with information structure?

Are predictions made under c.k. robust- in the sense of being approximamtely right- if there is "almost c.k."
And how should we define “almost c.k.” ?

Background: Consider a family of finite strategic form games with payoff functions that vary continuously with a parameter λ.
Then the set of NE is upper hemi-continuous in λ:

if σ^n is a NE for payoffs corresponding to λ^n, $\lambda^n \to \lambda$, $\sigma^n \to \sigma$, then σ is a NE for λ (see section 1.3.2 of FT)

So we don’t lose equilibria passing to a limit.
We know we can gain NE in the limit: the set of NE isn’t lower hemi-continuous.

An easy example:

	L	R
U	$1,1$	$0,0$
D	$0,0$	$0,0$

(D,R) is a NE in weakly dominated strategies.

if we look at a sequence of games where the payoff to (D,R) is $(-\varepsilon,-\varepsilon)$ no NE where players play D,R.

However, for generic payoff functions the set of NE is l.h.c in payoffs- this is discussed in FT section 12.1.2

These results are all stated in terms of changes to payoff functions, but they carry over immediately to changes in the prior distribution in the following way:

suppose there is a finite number of payoff matrices $u^1,\dots,u^L$ for finite strategy sets $S_1,\dots,S_I$

Finite state space Ω, common prior p, partitions H_i, and a map λ so that payoff functions in state ω are $u^{\lambda(\omega)}(.)$; the strategy spaces are now maps from H_i into S_i.

Since Ω is finite then this is a finite strategic form game. Now consider a sequence $p^n \to p$. Changing p just changes the payoff functions; so we know that once again the set of NE is uhc in p, not always lhc, but generically lhc.

Moreover if σ is a strict NE for p, then sigma is also a strict NE for p^n near p.

lhc can fail in the same way as in the previous example:

Prob $1-\varepsilon$ of
$$\begin{array}{cc} 1,1 & 0,0 \\ 0,0 & 0,0 \end{array}$$

Prob ε of
$$\begin{array}{cc} 1,1 & 1,-1 \\ -1,1 & -1,-1 \end{array}$$

trivial partitions - neither player sees the state, one info set each so the Bayesian game has matrix

$$\begin{array}{cc} 1,1 & \varepsilon,-\varepsilon \\ -\varepsilon,\varepsilon & -\varepsilon,-\varepsilon \end{array}$$

U dominates D, L dominate R
so none of the equil as ε goes to 0 look like the equil
D,R for $\varepsilon = 0$