8

Common Knowledge and Nash Equilibrium

> Where every man is Enemy to every man…the life of man is solitary, poore, nasty, brutish, and short.
>
> Thomas Hobbes

> In the case of any person whose judgment is really deserving of confidence, how has it become so? Because he has kept his mind open to criticism of his opinions …
>
> John Stuart Mill

This chapter applies the modal logic of knowledge developed in §4.1 and §5.10 to explore sufficient conditions for Nash equilibrium in two-player games (§8.1). We then expand the modal logic of knowledge to multiple agents and prove a remarkable theorem, due to Aumann (1976) that asserts that an event that is self-evident for each member of a group is common knowledge (§8.3).

This theorem is surprising because it appears to prove that individuals know the content of the minds of others with no explicit epistemological assumptions. We show in §8.4 that this theorem is the result of implicit epistemological assumptions involved in the construction of the standard semantic model of common knowledge, and when more plausible assumptions are employed, the theorem is no longer true.

Aumann's famous *agreement theorem* is the subject of section §8.7, where we show that the Aumann and Brandenburger (1995) conditions for Nash equilibrium in multi-player games is essentially an agreement theorem. Because there is no principle of Bayesian rationality that gives us the commonality of beliefs on which agreement depends, our analysis entails the demise of methodological individualism, a theme explored in section §8.8.

8.1 Conditions for Nash Equilibrium in Two-Player Games

Suppose that rational agents know one another's conjectures (§4.1), so that if $\phi_i^\omega(s_{-i}) > 0$ and $s_j \in S_j$ is player j's pure strategy in s_{-i}, then s_j is a best response to his conjecture ϕ_j^ω. We then have a genuine "equilibrium in conjectures," as now no agent has an incentive to change his pure strategy choice s_i, given the conjectures of the other players. We also have

THEOREM 8.1 *Let $\mathcal{G}$ be an epistemic game with Bayesian rational players, and suppose in state ω each player i knows the others' actions $\mathbf{s}_{-i}(\omega)$. Then $\mathbf{s}(\omega)$ is a Nash equilibrium.*

PROOF: To prove this theorem, which is due to Aumann xxand Brandenburger (1995), note that for each i, i knows the other players' actions at ω, so $\phi_i^{\omega}(s_{-i}) = 1$, which implies $\mathbf{s}_{-i}(\omega) = s_{-i}$ by K3, and i's Bayesian rationality at ω then implies $\mathbf{s}_i(\omega)$ is a best response to s_{-i}. ■

We say a Nash equilibrium in conjectures $(\phi_1^{\omega}, \dots, \phi_n^{\omega})$ occurs at ω if for each player i, $\mathbf{s}_i(\omega)$ is a best response to ϕ_i^{ω}, and for each i, $\phi_i^{\omega} \in \Delta^* S_{-i}$.We then have

THEOREM 8.2 *Suppose $\mathcal{G}$ is a two-player game, and at $\omega \in \Omega$, for $i = 1, 2$, $j \neq i$,*

1. *Each player knows the other is rational: i.e., $\forall \omega' \in \mathbf{P}_i\omega$, $\mathbf{s}_j(\omega')$ is a best response to $\phi_j^{\omega'}$;*
2. *Each player knows the other's beliefs; i.e., $\mathbf{P}_i\omega \subseteq \{\omega' \in \Omega | \phi_j^{\omega'} = \phi_j^{\omega}\}$.*

Then, the mixed strategy profile $(\sigma_1, \sigma_2) = (\phi_2^{\omega}, \phi_1^{\omega})$ is a Nash equilibrium in conjectures.

PROOF: To prove the theorem, which is due to Aumann xxand Brandenburger (1995) and Osborne and Rubinstein (1994), suppose s_1 has positive weight in $\sigma_1 = \phi_2^{\omega}$. Because $\phi_2^{\omega}(s_1) > 0$, there is some ω' such that $\omega' \in \mathbf{P}_2\omega$ and $s_1(\omega') = s_1$. By (1) s_1 is a best reply to $\phi_1^{\omega'}$, which is equal to ϕ_1^{ω} by (2). Thus s_1 is a best reply to $\sigma_2 = \phi_1^{\omega}$, and a parallel argument shows that s_2 is a best reply to σ_1, so (σ_1, σ_2) is a Nash equilibrium. ■

8.2 A Three-player Counterexample

	L	R
U	2,3,0	2,0,0
D	0,3,0	0,0,0

W

	L	R
U	0,0,0	0,2,0
D	3,0,0	3,2,0

E

Figure 8.1. Alice, Bob and Carole

Unfortunately, Theorem 8.2 does not extend to three or more players. For example Figure 8.1 shows a game where Alice chooses the row (U, D), Bob chooses the column (L, R), and Carole chooses the matrix (E, W) (the example is due to Osborne and Rubinstein, 1994:79). Note that every strategy of Carole's is a best response, because her payoff is identically zero. We assume there are seven states, so $\Omega = \{\omega_1, \ldots, \omega_7\}$, as depicted in Figure 8.2. States ω_1 and ω_7 represent Nash equilibria. There are also two sets of mixed strategy Nash equilibria. In the first, Alice plays D, Carole plays $2/5W + 3/5E$, and Bob plays anything (Carole's strategy is indeed specified by the condition that it gives Bob equal payoffs for all strategies), while in the second, Bob plays L, Carole plays $3/5W + 2/5E$, and Alice plays anything (this time, Carole's strategy is specified by the condition that it equalizes all Alice's payoffs).

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
P	32/95	16/95	8/95	4/95	2/95	1/95	32/95
s_1	U	D	D	D	D	D	D
s_2	L	L	L	L	L	L	R
s_3	W	E	W	E	W	E	E
$\mathcal{P}_A$	$\{\omega_1\}$	$\{\omega_2$	$\omega_3\}$	$\{\omega_4$	$\omega_5\}$	$\{\omega_6\}$	$\{\omega_7\}$
$\mathcal{P}_B$	$\{\omega_1$	$\omega_2\}$	$\{\omega_3$	$\omega_4\}$	$\{\omega_5$	$\omega_6\}$	$\{\omega_7\}$
$\mathcal{P}_C$	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_3\}$	$\{\omega_4\}$	$\{\omega_5\}$	$\{\omega_6\}$	$\{\omega_7\}$

Figure 8.2. Information Structure for Alice, Bob, and Carole game. Note that P is the probability of the state, s_i is i's choice in the corresponding state, and $\mathcal{P}_i$ is the knowledge partition for individual i.

Because there is a common prior (the 'P' row in Figure 8.2), and every state is in the corresponding cell of partition for each player (the last three rows in the figure), these are true knowledge partitions. Moreover, the posterior probabilities for the players are compatible with the knowledge operators for each player. For instance, in state ω_4, $\mathbf{P}_A\omega_4 = \{\omega_4, \omega_5\}$, and the conditional probability of ω_4, given $\mathbf{P}_A\omega_4$, is 2/3, and that of ω_5 is 1/3. Therefore, Alice's conjecture for Bob is $\phi_{AB}^{\omega_4} = L$, and for Carole is $\phi_{AC}^{\omega_4} = 2/3E + 1/3W$. Alice's move at ω_4, which is D, is therefore a best response, with payoff 2 as opposed the payoff of 2/3 from playing U against L and $2/3E + 1/3W$. Moreover, Alice knows that Carole is rational at ω_4 (trivially, because her payoff does not depend on her move). Alice knows Bob's beliefs at ω_4, because Bob could be either in

$\mathcal{P}_B$ partition cell $\{\omega_3, \omega_4\}$ or $\{\omega_5, \omega_6\}$, in both of which he believes Alice plays D and Carole plays $2/3W + 1/3E$. She also knows that Bob plays L in both cells, and Bob is rational because L pays off 2 against D and $2/3W + 1/3E$, as opposed to payoff 2/3 to playing R. Similarly, at ω_4, $\mathbf{P}_B\omega_4 = \{\omega_3, \omega_4\}$, so Bob knows that Alice is in either $\mathcal{P}_A$ partition cell $\{\omega_2, \omega_3\}$ or $\{\omega_4, \omega_5\}$, in both of which Alice knows that Bob plays L and Carole plays $2/3E+1/3W$. Thus, Bob knows Alice's beliefs and that Alice is rational in playing D. Similar reasoning shows that, Carole knows Alice and Bob's beliefs, and that they are rational at ω_4. Thus, all the conditions of the previous theorem are satisfied at ω_4, but of course, the conjectures at ω_4 do not form a Nash equilibrium, because $\phi_{AB}^{\omega_4} = L$ and $\phi_{BA}^{\omega_4} = D$ are not part of any Nash equilibrium of the game.

The reason Theorem 8.2 does not extend to this three player game is that Alice and Bob have different conjectures as to Carole's behavior, which is possible because Carole has more than one best response to Alice and Bob. They both know Carole is rational and they both know Carole believes $\phi_C^{\omega} = \{D, L\}$, for $\omega \in \{\omega_2, \dots, \omega_5\}$. However, these do not determine Carole's mixed strategy. Thus, mutual knowledge of rationality and beliefs is not sufficient to ensure that a Nash equilibrium will be played.

8.3 The Modal Logic of Common Knowledge

Suppose we have a set of n of agents, each of whom has a knowledge operator $\mathbf{K}_i$, $i = 1, \dots, n$. We say $E \subseteq \Omega$ is a *public event* if E is self-evident for all $i = 1, \dots, n$. By K1, Ω is a public event, and if E and F are public events, so is $E \cap F$, by K2a. Hence, for any $\omega \in \Omega$, there is a minimal public event $\mathbf{P}_*\omega$ containing ω; namely the intersection of all public events containing ω.

We can construct $\mathbf{P}_*\omega$ as follows. First, let

$$\mathbf{P}_*^1\omega = \bigcup_{j \in N} \mathbf{P}_j\omega, \tag{8.1}$$

which is the set of states that are possible for at least one agent at ω. Now, ω is possible for all players i from every state $\omega' \in \mathbf{P}_*^1\omega$, but an arbitrary $\omega' \in \mathbf{P}_*^1\omega$ is possible for some player i at ω, although not necessarily for all. So, $\mathbf{P}_*^1\omega$ may not be a public event. Thus we define

$$\mathbf{P}_*^2\omega = \bigcup\{\mathbf{P}_*^1\omega' | \omega' \in \mathbf{P}_*^1\omega\}, \tag{8.2}$$

which is the set of states that are possible for some agent at some state in $\mathbf{P}_*^1\omega$; i.e., this is the set of states that are possible for some agent from some state ω' that is possible for some (possibly other) agent at ω. Using similar reasoning, we see that any state in $\mathbf{P}_*^1$ is possible for any player i and any state $\omega' \in \mathbf{P}_*^2$, but there may be states in $\mathbf{P}_*^2\omega$ that are possible for one or more agents, but not all agents. In general, having defined $\mathbf{P}_*^i\omega$ for $i = 1, \ldots, k-1$, we define

$$\mathbf{P}_*^k\omega = \bigcup \{\mathbf{P}_*^1\omega' | \omega' \in \mathbf{P}_*^{k-1}\omega\}. \tag{8.3}$$

Finally, we define

$$\mathbf{P}_*\omega = \bigcup_{k=1}^{\infty} \mathbf{P}_*^k\omega. \tag{8.4}$$

This is the set of states ω' such that there is a sequence of states $\omega = \omega_1, \omega_2, \ldots, \omega_{k-1}, \omega_k = \omega'$ such that ω_{r+1} is possible for some agent at ω_r, for $r = 0, \ldots, r-1$. Of course, this is really a finite union, because Ω is a finite set. Therefore, for some k, $\mathbf{P}_*^k\omega = \mathbf{P}_*^{k+i}\omega$ for all $i \geq 1$.

We can show that $\mathbf{P}_*\omega$ is the minimal public event containing ω. First, $\mathbf{P}_*\omega$ is self-evident for each $i = 1, \ldots, n$, because for every $\omega' \in \mathbf{P}_*\omega$, $\omega' \in \mathbf{P}_*^k\omega$ for some integer $k \geq 1$, so $\mathbf{P}_i\omega' \subseteq \mathbf{P}_*^{k+1}\omega \subseteq \mathbf{P}_*\omega$. Hence $\mathbf{P}_*\omega$ is a public event containing ω. Now let E be any public event containing ω. Then, E must contain $\mathbf{P}_i\omega$ for all $i = 1, \ldots, n$, so $\mathbf{P}_*^1\omega \subseteq E$. Assume we have proven $\mathbf{P}_*^j\omega \subseteq E$ for $j = 1, \ldots, k$. Because $\mathbf{P}_*^k\omega \subseteq E$ and E is a public event, then $\mathbf{P}_*^{k+1}\omega = \mathbf{P}_*^1(\mathbf{P}_*^k\omega) \subseteq E$. Thus $\mathbf{P}_*\omega \subseteq E$.

The concept of a public even can be defined directly in terms of the agents' partitions $\mathcal{P}_1, \ldots, \mathcal{P}_n$. We say partition $\mathcal{P}$ is *coarser* than partition $\mathcal{Q}$ if every cell of $\mathcal{Q}$ lies in some cell of $\mathcal{P}$, and we say $\mathcal{P}$ is *finer* than $\mathcal{Q}$ if $\mathcal{Q}$ is coarser than $\mathcal{P}$. The public event partition $\mathcal{P}_*$ corresponding to $\mathbf{P}_*$ is then the finest common coarsening of the partitions $\mathcal{P}_1, \ldots, \mathcal{P}_n$ of the individual players.

To visualize these concepts, we return to the corn field analogy (§4.1). To coarsen a partition, simply remove one or more fence segments, and then to be tidy, repeatedly remove any fence segments that have either end unconnected to another segment. To refine (i.e., make finer) a partition, simply partition one or more of its cells. If the field has two partitions, visualize one with fence segments colored red and the other with segments colored blue. Where the fence segments intersect, let them share a common fence pole. Where a red and a blue fence segment separate the same corn stalks, including the fence segments surrounding the whole corn field, merge them into red and blue striped fence segments. The finest common coarsening of the two partitions is then the partition formed by removing all fence segments that are of only one color.

This visualization extends directly to the public event partition corresponding to the knowledge partitions in an n-player game. We give each player's fence partition a distinctive color, and we allow two or more agents to share fence segments by applying multiple colors to shared segments. We allow fence segments of different agents to pass through one another by placing a common fence pole at a point of intersection. Now, remove all fence segments that have fewer than n colors. What remains is the public event partition. Alternatively, the minimal public event $\mathbf{P}_*\omega$ containing state ω consists of the states that can be attained by walking from ω to any state in the field, provided one never climbs over a fence shared by all players.

Clearly the operator $\mathbf{P}_*$ satisfies P1. To show that it also satisfies P2, suppose $\omega' \in \mathbf{P}_*\omega$. Then, by construction, $\mathbf{P}_*\omega' \subseteq \mathbf{P}_*\omega$. To show that $\mathbf{P}_*\omega' = \mathbf{P}_*\omega$, note that $\omega' \in \mathbf{P}_*^k\omega$ for some k. Therefore, by construction, there is a sequence $\omega = \omega_1 = \ldots = \omega_k = \omega'$, such that $\omega_{j+1} \in \mathbf{P}_{i_j}\omega_j$ for some $i_j \in n$, for $j = 1, \ldots, k-1$. However, reversing the order of the sequence shows that $\omega \in \mathbf{P}_*\omega'$. Therefore $\mathbf{P}_*\omega = \mathbf{P}_*\omega'$. This proves that P2 holds, so $\mathbf{P}_*$ has all the properties of a possibilities operator.

It follows that $\mathbf{P}_*$ is a possibility operator. We define a *public event* operator $\mathbf{K}_*$ as the knowledge operator corresponding to the possibility operator $\mathbf{P}_*$, so $\mathbf{K}_*E = \{\omega | \mathbf{P}_*\omega \subseteq E\}$. We can then define an event E as a *public event* at $\omega \in \Omega$ if $\mathbf{P}_*\omega \subseteq E$. Thus, E is a public event if and only if E is self-evident to all players at each $\omega \in E$. Also, E is a public event if and only if E is the union of minimal public events of the form $\mathbf{P}_*\omega$. Moreover, K5 shows that if E is a public event, then at every $\omega \in E$ everyone knows that E is a public event at ω.

In the standard treatment of common knowledge (Lewis 1969, Aumann 1976), an event is common knowledge if everyone knows E, everyone knows that everyone knows E, and so on. It is easy to see a public event is always common knowledge, and conversely. For, suppose E is a public event. Then, for any $i, j, k = 1, \ldots, n$, $\mathbf{K}_iE = E$, $\mathbf{K}_j\mathbf{K}_iE = \mathbf{K}_jE = E$, $\mathbf{K}_k\mathbf{K}_j\mathbf{K}_iE = \mathbf{K}_kE = E$, and so on. Thus all events of the form $\mathbf{K}_k\mathbf{K}_j \ldots \mathbf{K}_iE$ are self-evident for k, so E is common knowledge. Conversely, suppose that for any sequence $i, j, \ldots, k = 1, \ldots, n$,

$\mathbf{K}_i\mathbf{K}_j \ldots \mathbf{K}_k E \subseteq E$. Then, for any $\omega \in E$, because $\mathbf{P}_i\omega \subseteq E$, we have $\mathbf{P}^1_*\omega \subseteq E$, where $\mathbf{P}^1_*$ is defined in (8.1). We also have $\mathbf{K}_i\mathbf{P}^1_*\omega \subseteq E$, because $\mathbf{K}_i\mathbf{K}_j E \subseteq E$ for $i, j = 1, \ldots, n$, so $\mathbf{P}^2_*\omega \subseteq E$ from (8.2). From (8.3), we now see that $\mathbf{P}^k_*\omega \subseteq E$ for all k, so $\mathbf{P}_*\omega \subseteq E$. Therefore E is the union of public events, and hence is a public event.

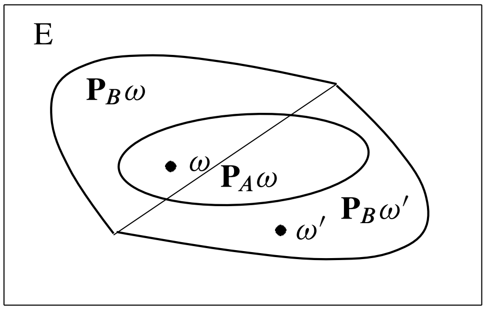

Figure 8.3. At ω, Bob knows that Alice knows that E.

Figure 8.3 shows the situation where Alice knows E at ω, because her minimal self-evident event $\mathbf{P}_A\omega$ at ω lies within E. Moreover $\mathbf{P}_A\omega$ intersects two of Bob's minimal self-evident events, $\mathbf{P}_B\omega$ and $\mathbf{P}_B\omega'$. Because both of $\mathbf{P}_B\omega$ and $\mathbf{P}_B\omega'$ lie within E, Bob knows that Alice knows that E at ω (and at every other state in $\mathbf{P}_A\omega$).

8.4 The Commonality of Knowledge

We have defined a public event as an event that is self-evident to all players. We then showed that an event E is public if and only if, it is common knowledge. It appears, then, that at a public event, there is a perfect *commonality of knowledge*: players know a great deal about what other players know. Where does this knowledge come from? The answer is that we have tacitly assumed that the way each individual partitions Ω is "known" to all, not it the formal sense of a knowledge operator, but rather in the sense that an expression of the form $\mathbf{K}_i\mathbf{K}_j E$ makes sense, and means "i knows that j knows that E." Formally, to say that i knows that j knows E at ω means that at every state $\omega' \in \mathbf{P}_j\omega$, $\mathbf{P}_i\omega' \subseteq E$. But, i knows that this is the case only if he "knows" $\mathbf{P}_j\omega$, which allows him to test $\mathbf{K}_i\omega' \subseteq E$ for each $\omega' \in \mathbf{P}_j\omega$.

For example, suppose Alice, Bob, and Carole meet yearly on a certain date at a certain time to play a game $\mathcal{G}$. Suppose, by chance, all three happened to be in Dallas, Texas the day before, and although they did not see each other, each witness the same highly unusual event x. We define the universe $\Omega = \{\omega, \omega'\}$, where the unusual even occurs in ω but not in ω'. Then, $\mathbf{P}_A\omega = \mathbf{P}_B\omega = \mathbf{P}_C\omega = \{\omega\}$, and hence $\mathbf{K}_A\omega = \mathbf{K}_B\omega = \mathbf{K}_C\omega = \{\omega\}$. Thus ω is self-evident to all three individuals, and hence ω is a public event. Therefore at ω, Alice knows that Bob knows that Carole knows ω, and so on. But, of course, this is not the case. Indeed, none of the three individuals is aware that the others know the event x.

The problem is that we have misspecified the universe. Suppose an event ω is a four-vector, the first entry of which is either x or $\neg x$ (meaning "not x"), and the other three entries are "true" or "false," depending on whether Alice, Bob, and Carole, respectively, knows or does not know whether x occurred. The universe Ω now has sixteen distinct states, and the state ω that actually occurred is $\omega = [x,\text{true},\text{true},\text{true}]$. However, now $\mathbf{P}_A\omega = \{\omega' \in \Omega | \omega'[1] = x \wedge \omega'[2] = \text{true}\}$. Therefore, the state ω is now *not* self-evident for Alice. Indeed, the smallest self-evident event $\mathbf{P}_A\omega$ for Alice at ω in this case is Ω itself!

This line of reasoning reveals a central lacuna in epistemic game theory: its semantic model of common knowledge assumes too much. Economists have been misled by the elegant theorem that says mutual self-evidence implies common knowledge into believing the axioms of rational choice imply something substantive concerning the commonality of knowledge across agents. They do not. Indeed, there is no formal principle specifying conditions under which distinct individuals will attribute the same truth-value to a proposition p with empirical content (we can assume rational agents will all agree on mathematical and logical tautologies), or will have a mental representation of the fact that others attribute truth-value to p. We address this below by sketching the attributes of what we have termed *mutually accessible* events (§7.8).

8.5 The Tactful Ladies

While walking in the garden, Alice, Bonnie and Carole encountered a violent thunderstorm and are obliged to duck hastily into a restaurant for tea. Carole notices that Alice and Bonnie have dirty foreheads, although each is unaware of this fact. Carole is too tactful to mention this embarrassing situation, which would surely lead them to blush, but she observes that, like herself, each of the two ladies knows that someone has a dirty forehead but is also too tactful to mention this fact. The thought occurs to Carole that she also might have a dirty forehead, but there are no mirrors or other detection devices handy that might help resolve her uncertainty.

At this point, a little boy walks by the three young ladies' table and exclaims "I see a dirty forehead!" After a few moments of awkward silence, Carole realizes that she has a dirty forehead, and blushes.

How is this feat of logical deduction possible? Certainly it is mutually known among the ladies that at least one of them had a dirty forehead, so the little boy did not inform any of them of this fact. Moreover, each lady could see that the other ladies each saw at least one dirty forehead, so it is mutually known that each lady knew what the little boy said before he said it. However, the little boy's remark does inform each lady that they all know that they all know that one of them has a dirty forehead. This is something than none of the ladies knew before the little boy's announcement. For instance, Alice and Bonnie each knows she might not have a dirty forehead, so Alice knows that Bonnie might believe that Carole sees two clean foreheads, in which case Alice and Bonnie know that Carole might not know that there is at least one dirty forehead. Following the little boy's announcement, however, and assuming the other ladies are logical thinkers (which they must be if they are Bayesian decision-makers), Carole's inference concerning the state of her forehead is unavoidable.

To see why, suppose Carole did not have a dirty forehead. Carole then knows that Alice sees one dirty forehead (Bonnie's), so Alice learns nothing from the little boy's remark. But, Carole knows that Bonnie sees that Carole's forehead is not dirty, so if Bonnie's forehead were not dirty, then Alice would have seen two clean foreheads, and the little boy's remark would imply that Alice would know that she was the unfortunate possessor of a dirty forehead. Because Alice did not blush, Carole knows that Bonnie would conclude that she must have a dirty forehead, and would have blushed. Because Bonnie did no such thing, Carole knows that her assumption that she has a clean forehead is false.

To analyze this problem formally, suppose Ω consists of eight states of the form $\omega = xyz$, where $x, y, z \in \{d, c\}$ are the states of Alice, Bonnie, and Carole, respectively and where d and c stand for "dirty forehead" and "clean forehead," respectively. Thus, for instance $\omega = ccd$ is the state of the world where Carole has a dirty forehead but Alice and Bonnie both have clean foreheads. When Carole sits down to tea, she knows $E_C = \{ddc, ddd\}$, meaning she sees that Alice and Bonnie have dirty foreheads, but her own forehead could be either clean or dirty. Similarly, Alice knows

$E_A = \{cdd, ddd\}$ and Bonnie knows $E_B = \{dcd, ddd\}$. Clearly, no lady knows her own state. What does Bonnie know about Alice's knowledge? Because Bonnie does not know the state of her own forehead, she knows that Alice knows the event "Carole has a dirty forehead," which is $E_{BA} = \{cdd, ddd, ccd, dcd\}$. Similarly, Carole knows that Bonnie knows that Alice knows $E_{CBA} = \{cdd, ddd, ccd, dcd, cdc, ddc, ccc, dcc\} = \Omega$. Assuming Carole has a clean forehead, she knows that Bonnie knows that Alice knows $E'_{CBA} = \{cdc, ddc, dcc, ccc\}$. After the little boy's announcement, Carole would then know that Bonnie knows that Alice knows $E''_{CBA} = \{cdc, ddc, dcc\}$, so if Bonnie did not have a dirty forehead, she would know that Alice knows $E''_{BA} = \{dcc\}$, so Bonnie would conclude that Alice would blush. Thus, Bonnie's assumption that she has a clean forehead would be incorrect, and she would blush. Because Bonnie does not blush, Carole knows that her assumption that she has a clean forehead is incorrect.

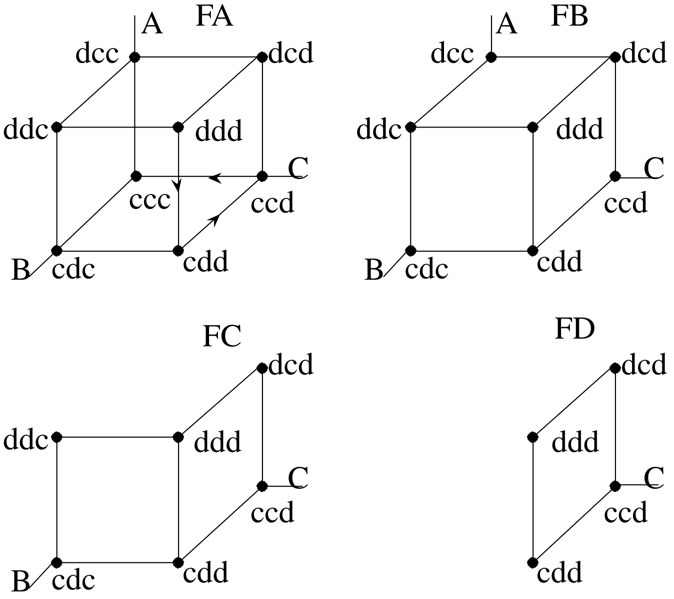

Figure 8.4. The Three Ladies Problem

There is an instructive visual way to approach the problem of the tactful ladies, due to Fagin, Halpern, Moses and Vardi (1995) and illustrated in figure 8.4. Think of each of the ladies owning one of the three axes in this figure, each corner of the cube representing one of the eight states of the world. The endpoints of lines parallel to an axis represent minimal self-evident events for the lady owning that axis; i.e., the lady in question cannot determine whether her own forehead is dirty.

Because the endpoints of every line segment is a minimal self-evident event for one of the ladies, a node is reachable from another provided there is some path along the lines of the graph, connecting the first to the second. What, for instance, does it mean that ccc is reachable from ddd along the arrows in pane FA of the figure? First, at ddd, Alice believes cdd is possible, at cdd, B believes ccd is possible, and at ccd, C believes that ccc is possible. In other words, at ddd, Alice believes that it is possible that B believes that it is possible that C believes that ccc might be the true state. Indeed, it is easy to see that any sequence of moves around the cube corresponds to some statement of the form x believes it is possible that y believes it is possible that..., and so on. We define an event $E \subseteq \Omega$ as a public event, or common knowledge, if every state $\omega \in E$ is reachable from every other in this manner. Clearly, the only public event is Ω itself.

When the little boy announces b (someone has a dirty forehead), and assuming this statement is taken as truthful, then the three ladies then all know that ccc cannot occur, so we can delete all the paths from some state to ccc. The result is shown in pane FB of the figure. Now, if dcc were the state, Alice would know she has a dirty forehead, and because she apparently does not know this, we can delete the lines terminating in dcc, leading to pane FC in the figure. Now, at ddc or cdc, Bonnie would know she has a dirty forehead, so we can delete the lines connecting to these two nodes. This leaves the nodes depicted in pane FD. Clearly, Carole knows at this event that she has a dirty forehead, but Alice and Bonnie do not.

8.6 The Tactful Ladies and the Commonality of Knowledge

The Three Tactful Ladies problem involves many unstated epistemological assertions going far beyond the common knowledge of rationality involved in the conclusion that Carole knows the state of her forehead. Let us see exactly what they are.

Let x_i be the condition that i has a dirty forehead, and let k_i be the knowledge operator for i, where $i = A, B, C$, standing for Alice, Bonnie, and Carole, respectively. When we write i, we mean any $i = A, B, C$, and when we write i, j, we mean any $i, j = A, B, C$ with $j \neq i$, and when we write i, j, m we mean $i, j, m = A, B, C$ and $i \neq j \neq m \neq i$. Let y_i be the condition that i blushes. The six symbols x_i and y_i represent the possible states