Common Knowledge, Coherent Uncertainties and Consensus

Yakov Ben-Haim
Faculty of Mechanical Engineering
Technion — Israel Institute of Technology
Haifa 32000 Israel
yakov@aluf.technion.ac.il

Abstract

In this paper we study info-gap analogs of the classical game-theoretic concepts of information- and knowledge-functions, common knowledge and consensus. Our main results are that knowledge is constricted as info-gap-uncertainty grows (theorem 1), common knowledge is limited by the presence of info-gap uncertainty (theorem 2), and that common knowledge is related to consensus via the info-gap coherence functions (theorem 3). We discuss several examples, including the prisoners' dilemma, principal-agent contract negotiations, search and evasion strategies, and teamwork and the need for costly transfer of information.

Contents

papers adversity ccc.tex 12.12.01
Technion Report ETR-2001-03

1 Introduction

Aumann's seminal paper of 1976 established the critical role of common knowledge in the achievement of consensus. No less importantly, as Aumann himself notes, the result derives from an "appropriate framework", and thus establishes the fruitfulness of the concepts of partition and knowledge for the analysis of agreement under disparate information.

In this paper we develop an information-gap framework for studying the role of common knowledge in consensus among decision makers who operate with different information. We begin with a very brief summary of the axioms and structure of info-gap models of uncertainty (section 2). Then, in section 3, we explain how an info-gap model can be interpreted in a way which is similar to, though different from, a partitional information function. We derive a knowledge function in section 4 based on info-gap information functions, and compare it with the classical knowledge function based on a partitional information function. Many though not all of the familiar properties of classical knowledge functions are preserved. An important difference is that the info-gap knowledge function entails an explicit dimension of info-gap uncertainty, and theorem 1 shows that knowledge is constricted as the info-gap grows. Section 5 compares the info-gap knowledge function with the belief function studied by Monderer and Samet (1989).

In section 6 we define common knowledge in the info-gap context. Our most important result in this section, theorem 2, shows that common knowledge is limited by info-gap uncertainty. That is, common-knowledge sets become empty beyond some level of iteration unless the info-gap is zero. We apply this to a discussion of the prisoners' dilemma.

In section 7 we employ the concept of info-gap coherence functions, which assess the degree of similarity of distinct info-gap models of uncertainty. The coherence functions are important as an indication of agreement on preferences among decision makers with different information. Complete coherence guarantees consensus. We establish several relations between knowledge- and common-knowledge-functions and the coherence functions. We show in theorem 3 that complete coherence is equivalent to each agent being able to deduce common knowledge from self knowledge.

In section 8 we consider an example of contract negotiation between an employer and a potential employee. We show how the disparity between the agents' knowledge functions, and hence the loss of common knowledge, increases with the incoherence of their info-gap models. In section 9 we examine a general two-competitor situation such as a search-and-evasion problem. We show how each agent can use the info-gap self-knowledge functions strategically by identifying situations in which the other party cannot know about one's own knowledge. In section 10 we study a reverse situation in which information-transfer is needed for collaboration between design teams. Analysis of the knowledge functions facilitates efficient formulation of the costly transfer of information.

All proofs appear in the appendix.

2 Info-gap Models of Uncertainty

Our quantification of uncertainty is based on non-probabilistic information-gap models (Ben-Haim, 2001). The info-gap intuition is that uncertainty is a disparity between what the decision maker knows and what could be known. An info-gap model is a family of nested sets. Each set corresponds to a particular degree of uncertainty, according to its level of nesting. Each element in a set represents a possible realization of the uncertain event. There are no measure functions in an info-gap model of uncertainty.

The theory of info-gap uncertainty provides a quantitative model for Knight's concept of "true uncertainty" for which "there is no objective measure of the probability", as opposed to risk which is probabilistically measurable (Knight, 1921, pp.46, 120, 231–232). Further discussion of the relation between Knight's conception and info-gap theory is found in (Ben-Haim, 2001, section 12.5). Similarly, Shackle's "non-distributional uncertainty variable" bears some similarity to info-gap analysis

(Shackle, 1972, p.23). Likewise, Kyburg recognized the possibility of a "decision theory that is based on some non-probabilistic measure of uncertainty." (Kyburg, 1990, p.1094).

Uncertain quantities are vectors or vector functions. Uncertainty is expressed at two levels by info-gap models. For fixed α the set $\mathcal{U}(\alpha, \widetilde{x})$ represents a degree of variability of the uncertain quantity x around the centerpoint $\widetilde{x}$. The greater the value of α, the greater the range of possible variation, so α is called the *uncertainty parameter* and expresses the information gap between what is known ($\widetilde{x}$ and the structure of the sets) and what needs to be known for an ideal solution (the exact value of x). The value of α is usually unknown, which constitutes the second level of uncertainty: the horizon of uncertain variation is unbounded.

Let $\Re$ denote the non-negative real numbers and let Ω be a Banach space in which the uncertain quantities x are defined. An info-gap model $\mathcal{U}(\alpha, \widetilde{x})$ is a map from $\Re \times \Omega$ into the power set of Ω. Info-gap models obey four axioms. *Nesting:* $\mathcal{U}(\alpha, \widetilde{x}) \subseteq \mathcal{U}(\alpha', \widetilde{x})$ if $\alpha \leq \alpha'$. *Contraction:* $\mathcal{U}(0, 0)$ is the singleton set $\{0\}$. *Translation:* $\mathcal{U}(\alpha, \widetilde{x})$ is obtained by shifting $\mathcal{U}(\alpha, 0)$ from the origin to $\widetilde{x}$: $\mathcal{U}(\alpha, \widetilde{x}) = \mathcal{U}(\alpha, 0) + \widetilde{x}$. *Linear expansion:* info-gap models centered at the origin expand linearly: $\mathcal{U}(\alpha', 0) = \frac{\alpha'}{\alpha}\mathcal{U}(\alpha, 0)$ for all α, $\alpha' > 0$. For more discussion of these axioms see Ben-Haim (1999).

3 Information Functions

In this section we will explain the sense in which an info-gap model can be understood as an 'information function' in analogy to the classical results reviewed by Osborne and Rubinstein (1994, pp.67–70). In section 4 we will extend the discussion to an info-gap knowledge function.

Let Ω denote the set of all events or states: the universe of discourse. Ω need not be a Banach space, nor even an infinite set. A classical **information function** P maps every element $\omega \in \Omega$ to a nonempty subset $P(\omega)$ of Ω. The interpretation of the information function $P(\omega)$ "is that when the state is ω the decision maker knows only that the state is in the set $P(\omega)$. That is, he considers it possible that the true state could be any state in $P(\omega)$ but not any state outside $P(\omega)$." (Osborne and Rubinstein 1994, pp.67–68). See also (Aumann 1976, p.594).

$P(\omega)$ is set-theoretic, non-probabilistic, and suggestive of info-gap ideas. However, $P(\omega)$ specifies a precise boundary between the known and the unknown. This is softened in the info-gap approach without introducing distributional modelling. From the epistemic point of view, sharp and immutable assertions are not always compatible with severely deficient information. From the perspective of objective or ontological uncertainty, the real world may simply not display the sort of crisp and clear delineation which is entailed in a partition such as $P(\omega)$.

In this section we will explore an analogy between $P(\omega)$ and the info-gap model $\mathcal{U}(\alpha, \omega)$, $\alpha \geq 0$. We henceforth consider the universe of discourse Ω to be a Banach space. An interpretation of $\mathcal{U}(\alpha, \omega)$ as an information function is that the decision maker knows that, up to info-gap α, any state in $\mathcal{U}(\alpha, \omega)$ is possible. The decision maker's epistemic condition is the family of nested sets $\mathcal{U}(\alpha, \omega)$, $\alpha \geq 0$. The decision maker knows that the state of the world is constrained to some of the sets in this family, but he does not know which sets. What he *does* know is ω and the structure of the family of sets, but he *does not* know the values of α which capture the true state of the world, nor what that true state actually is. The decision maker suffers two levels of epistemic limitation: (1) for any given α he does not know what (if any) element of $\mathcal{U}(\alpha, \omega)$ holds, and (2) regarding α he knows only that $\alpha \geq 0$. Discussion of some further epistemological aspects of info-gap uncertainty is found in Ben-Haim (2001, chapter 12).

Classical information functions are usually assumed to obey two conditions:

$$P1 \qquad \omega \in P(\omega) \text{ for every } \omega \in \Omega \tag{1}$$

$$P2 \qquad \text{If } \omega' \in P(\omega) \text{ then } P(\omega') = P(\omega) \tag{2}$$

It is clear from the axioms of info-gap models that $\mathcal{U}(\alpha, \omega)$ obeys the analog of $P1$ but does not have a property analogous to $P2$. The info-gap analog of $P1$ is:

Lemma 1

$$\omega \in \mathcal{U}(\alpha, \omega) \tag{3}$$

Now consider $P2$. Many 'garden variety' info-gap models show that:

$$\omega' \in \mathcal{U}(\alpha, \omega) \quad \text{does not imply} \quad \mathcal{U}(\alpha, \omega') = \mathcal{U}(\alpha, \omega) \tag{4}$$

which, were the implication valid, would be the info-gap analog of $P2$. For instance, consider an ellipsoidal info-gap model:

$$\mathcal{U}(\alpha, \omega) = \left\{ \nu = \omega + x : \ x^T V x \le \alpha^2 \right\}, \quad \alpha \ge 0 \tag{5}$$

where V is a real, symmetric, positive definite matrix. $\mathcal{U}(\alpha, \omega)$, $\alpha \ge 0$, is the family of ellipsoids centered at ω whose shape is determined by V. If ω' lies on the boundary of one of these ellipsoids, then $\mathcal{U}(\alpha, \omega')$, $\alpha \ge 0$, is the family of ellipsoids of shape V centered at ω', which is a different family of sets from $\mathcal{U}(\alpha, \omega)$, $\alpha \ge 0$. Clearly $\omega' \in \mathcal{U}(\alpha, \omega)$ but $\mathcal{U}(\alpha, \omega') \ne \mathcal{U}(\alpha, \omega)$.

A partition of a set A is a collection of disjoint subsets of A whose union equals A. Osborne and Rubinstein show that an information function $P(\omega)$ is a partition of Ω if and only if P obeys conditions $P1$ and $P2$. An info-gap model — a family of nested sets $\mathcal{U}(\alpha, \omega)$, $\alpha \ge 0$ — is not a partition, though it is nonetheless amenable to a set-theoretic interpretation as a decision maker's information: $\mathcal{U}(\alpha, \omega)$ is what the decision maker knows about the universe of possibilities, up to info-gap α.

$P2$ is a particularly important property of classical information functions. Osborne and Rubinstein explain that

> P2 says that the decision maker uses the consistency or inconsistency of states with his information to make inferences about the state. Suppose, contrary to P2, that $\omega' \in P(\omega)$ and there is a state $\omega'' \in P(\omega')$ with $\omega'' \notin P(\omega)$. Then if the state is ω the decision maker can argue that since ω'' is inconsistent with his information the true state cannot be ω'. (Osborne and Rubinstein 1994, p.68).

Let's examine how the analogous argument would proceed in the info-gap context, in light of the fact that the analog of $P2$ does not hold.

Our info-gap model is the family of nested sets $\mathcal{U}(\alpha, \omega)$, $\alpha \ge 0$. Suppose that:

$$\omega' \in \mathcal{U}(\alpha, \omega) \tag{6}$$

which means that ω' is consistent with our information, up to uncertainty α. Now suppose there is an event ω'' such that:

$$\omega'' \in \mathcal{U}(\alpha, \omega') \ \text{ and } \ \omega'' \notin \mathcal{U}(\alpha, \omega) \tag{7}$$

That is, ω'' is consistent with ω' up to α, but inconsistent with ω up to α. Nonetheless, there is an $\alpha'' > \alpha$ for which:

$$\omega'' \in \mathcal{U}(\alpha'', \omega) \tag{8}$$

In other words, ω'' is consistent with our information up to uncertainty α''. The decision maker can conclude that ω'' is, in some sense, less consistent with our information than ω'. This may impugn ω' in some sense, (since ω' *is* consistent with our information up to α) but ω' cannot be excluded from the realm of the possible. Information expressed by an info-gap model is less informative, more flexible, and less committal than classical information functions.

4 Knowledge Functions

In this section we will define an info-gap knowledge function which is analogous to the classical results discussed by Osborne and Rubinstein (1994, pp.67–70).

For any event E, a subset of Ω, a classical information function P generates a **knowledge function** defined as:

$$K(E) = \{\omega \in \Omega : \; P(\omega) \subseteq E\} \tag{9}$$

Osborne and Rubinstein (1994, p.69) explain that "the set $K(E)$ is the set of all states in which the decision maker knows E." Regardless of whether or not P obeys $P1$ and $P2$, the knowledge function has the following properties: $K(\Omega) = \Omega$; $E \subseteq F$ implies that $K(E) \subseteq K(F)$; and $K(E) \cap K(F) = K(E \cap F)$.

We can define an **info-gap knowledge function** in a very similar way with respect to an info-gap model:

$$K_\alpha(E) = \{\omega : \; \mathcal{U}(\alpha, \omega) \subseteq E\} \tag{10}$$

$K_\alpha(E)$ is the set of states ω in which the decision maker knows, up to uncertainty α, that event E holds. Stated differently, $K_\alpha(E)$ is the set of states which, up to uncertainty α, are consistent with E.

In discussing info-gap knowledge functions we define the domain of discourse Ω as the Banach space in which the info-gap model is defined. It is then readily shown that $K_\alpha(E)$ obeys the analogs of the basic properties of classical knowledge functions:

Lemma 2 *The knowledge function $K_\alpha(E)$ has the properties:*

$$K1 \qquad K_\alpha(\Omega) = \Omega \tag{11}$$

$$K2 \qquad E \subseteq F \;\text{ implies that }\; K_\alpha(E) \subseteq K_\alpha(F) \tag{12}$$

$$K3 \qquad K_\alpha(E) \cap K_\alpha(F) = K_\alpha(E \cap F) \tag{13}$$

The following proposition is an additional immediate consequence of the nesting axiom of info-gap models:

Theorem 1 *Knowledge is constricted as uncertainty grows:*

$$\alpha < \alpha' \;\text{ implies that }\; K_{\alpha'}(E) \subseteq K_\alpha(E) \tag{14}$$

An info-gap model, $\mathcal{U}(\alpha, \omega)$, $\alpha \geq 0$, is a family of nested sets, where α is the uncertainty parameter. In theorem 1, α' is a greater horizon of uncertainty than α. The result indicates that the set of states in which we can know E is smaller at the greater horizon of uncertainty.

Theorem 1 is to be distinguished from the nesting property (axiom 1) of info-gap models:

$$\alpha < \alpha' \;\text{ implies that }\; \mathcal{U}(\alpha, \omega) \subseteq \mathcal{U}(\alpha', \omega) \tag{15}$$

An info-gap model is the analog of the classical partitional information function. The intuition in relation (15) is similar to the idea expressed by Fudenberg and Tirole (1991, p.543) that "more precise information corresponds to knowing a *smaller* set: Knowledge here is the ability to rule out some of the states that were possible *ex ante*." Regarding relation (15): the information at info-gap α (that is, $\mathcal{U}(\alpha, \omega)$) is "more precise" than the information at info-gap α' (namely $\mathcal{U}(\alpha', \omega)$) since more states can be ruled out at α than at α'. This greater precision is manifested in relation (14) by the fact that at α we can assert the truth of E in a wider class of states than at α'.

We see that $K(E)$ and $K_\alpha(E)$ have similar (though not identical) interpretations as knowledge functions, and they both obey properties $K1$–$K3$ or their analogs without invoking property $P1$

or $P2$. Likewise we recognize that $P(\omega)$ and $\mathcal{U}(\alpha,\omega)$ bear analogous interpretations as information functions.

Osborne and Rubinstein show that, if the information function P satisfies property $P1$, then the knowledge function obeys: $K(E) \subseteq E$. The analogous relation holds for $K_\alpha(E)$:

Lemma 3

$$K_\alpha(E) \subseteq E \tag{16}$$

If the classical information function $P(\omega)$ satisfies both $P1$ and $P2$ then its knowledge function obeys:

$$K(E) \subseteq K(K(E)) \tag{17}$$

$$\Omega - K(E) \subseteq K(\Omega - K(E)) \tag{18}$$

Osborne and Rubinstein note that in fact equality holds in both of these relations.

Inclusion (17) is interpreted by Osborne and Rubinstein to mean that "if the decision maker knows E then he knows that he knows E." (1994, p.70). In light of the fact that relation (17) is in fact an equality, its interpretation is also that 'knowledge of E' and 'knowledge of knowledge of E' are the same event. Whether this is reasonable depends upon, among other things, what exactly we mean by 'knowledge'. One sense of the term is "familiarity gained by experience" (Flexner, 1980; Oxford English Dictionary, 1999). Who has not had the delightfully surprising experience of discovering that one has an unanticipated ability, like for driving on ice or for installing complex computer programs. In such cases knowing (how to drive on ice) is distinct from knowing that one knows (how to drive on ice). The point is that (17) is not necessarily a requirement for a meaningful knowledge function.

Example 1 A simple example will illustrate that the info-gap analog of relation (17) does not necessarily hold. Let the domain of discourse Ω be the set of real numbers and consider the info-gap model $\mathcal{U}(\alpha,\omega)$ defined as the family of closed intervals centered at ω:

$$\mathcal{U}(\alpha,\omega) = \{v: \ |v-\omega| \le \alpha\}, \quad \alpha \ge 0 \tag{19}$$

Let E be any subset of Ω, so the knowledge function $K_\alpha(E)$ is the set of centerpoints of intervals of width 2α contained in E. In particular, for $E = [0,1]$:

$$K_\alpha(E) = \{\omega: \ \mathcal{U}(\alpha,\omega) \subseteq E\} \tag{20}$$

$$= \{\omega: \ [\omega-\alpha, \ \omega+\alpha] \subseteq [0,1]\} \tag{21}$$

$$= [\alpha, 1-\alpha] \tag{22}$$

which is empty unless $\alpha \le 1/2$.

In similar fashion we find:

$$K_\alpha(K_\alpha(E)) = \{\omega: \ \mathcal{U}(\alpha,\omega) \subseteq K_\alpha(E)\} \tag{23}$$

$$= \{\omega: \ [\omega-\alpha, \ \omega+\alpha] \subseteq [\alpha, 1-\alpha]\} \tag{24}$$

$$= [2\alpha, 1-2\alpha] \tag{25}$$

which is empty unless $\alpha \le 1/4$. We see that $K_\alpha(K_\alpha(E))$ is a proper subset of $K_\alpha(E)$ unless they are both empty, contradicting (17). In summary, in this example, 'knowing that we know E' entails 'knowing E', but not the converse. ∎

In fact, lemma 3 immediately implies that 'knowledge of knowledge of E' entails 'knowledge of E': $K_\alpha(K_\alpha(E)) \subseteq K_\alpha(E)$.

Relation (18) is interpreted by Osborne and Rubinstein (1994, p.70) to mean "that the decision maker is aware of what he does not know: if he does not know E then he knows that he does not know

E." A decision maker with this level of self-awareness is to be admired; awareness of ones ignorance is often purchased only at enormous cost. The following example illustrates that the info-gap analog of (18) need not hold.

Example 2 Continuing example 1, we find from eq.(22) that:

$$\Omega - K_\alpha(E) = \begin{cases} (-\infty, \alpha) \cup (1-\alpha, \infty) & \alpha \le \frac{1}{2} \\ (-\infty, \infty) & \alpha > \frac{1}{2} \end{cases} \tag{26}$$

From this we find that:

$$K_\alpha(\Omega - K_\alpha(E)) = \begin{cases} (-\infty, 0) \cup (1, \infty) & \alpha \le \frac{1}{2} \\ (-\infty, \infty) & \alpha > \frac{1}{2} \end{cases} \tag{27}$$

We see that, in this example:

$$\Omega - K_\alpha(E) \not\subseteq K_\alpha(\Omega - K_\alpha(E)) \tag{28}$$

for $0 < \alpha < 1/2$, which shows that the info-gap analog of relation (18) need not hold. ■

5 Knowledge and Belief

In this section we briefly point out a parallel between the info-gap knowledge function and a probabilistic conception of belief.

For an agent whose information structure is a partition $P(\omega)$ with probability measure μ, Monderer and Samet (1989) denote the event 'the agent p-believes E' by:

$$B^p(E) = \{\omega : \ \mu(E|P(\omega)) \ge p\} \tag{29}$$

Thus $B^p(E)$ is the set of states of the world ω in which the agent's information $P(\omega)$ indicates that the probability of E is no less than p.

Now consider the info-gap knowledge function based on the info-gap model $\mathcal{U}(\alpha, \omega)$. The agent's info-gap knowledge function, $K_\alpha(E)$ in eq.(10), is the set of states of the world ω in which the agent can conclude, at ambient uncertainty no greater than α, that E holds. If the agent's info-gap is no greater than α, then $K_\alpha(E)$ is the set of states in which he can infer E. Motivated by the interpretation of $B^p(E)$ as the event 'p-belief in E', we see that $K_\alpha(E)$ can be interpreted as the event 'α-belief in E'. The difference of course is that p-belief entails probabilistic knowledge, the measure function μ, while info-gap-belief entails the non-probabilistic info-gap model $\mathcal{U}$.

One way to further establish the parallel between p- and α-belief is to note that they display the same inverse-nesting:

$$p < p' \quad \text{implies that} \quad B^{p'}(E) \subseteq B^p(E) \tag{30}$$

$$\alpha < \alpha' \quad \text{implies that} \quad K_{\alpha'}(E) \subseteq K_\alpha(E) \tag{31}$$

Relation (30) states that one's p-belief diminishes as one demands greater certainty for the belief. That is, the set of states in which one can be very confident (p'-confident) that E holds, is smaller than the set of states in which one's confidence is only p. Likewise, relation (31) (which is theorem 1) states that one's α-belief diminishes as one demands greater assertability for the belief. That is, the set of states in which one can assert E even in the presence of a large info-gap, α', is smaller than the set of states in which one can assert E when one's uncertainty is only α.

6 Common Knowledge

In this section we discuss a direct info-gap analog of the classical idea of common knowledge.

Let the knowledge functions of two agents be $K_{1,\alpha}(E)$ and $K_{2,\alpha}(E)$, based on info-gap models $\mathcal{U}_1(\alpha,\omega)$ and $\mathcal{U}_2(\alpha,\omega)$, respectively. The n-fold iterated knowledge function of agent i is, for $j = 3-i$ and for $n \geq 1$:

$$K^n_{i,\alpha}(E) = K_{i,\alpha}(K^{n-1}_{j,\alpha}(E)) \tag{32}$$

where we define $K^0_{i,\alpha}(E) = E$. The interpretation is that $K_{i,\alpha}(E)$ is the set of states in which i can infer that, up to info-gap α, E holds. Thus $K^2_{i,\alpha}(E) = K_{i,\alpha}(K_{j,\alpha}(E))$ is the set of states in which i infers (up to α) that j infers (up to α) that E holds. In the set $K^3_{i,\alpha}(E) = K_{i,\alpha}(K_{j,\alpha}(K_{i,\alpha}(E)))$, i infers (up to α) that j infers (up to α) that i infers (up to α) that E holds. And on it goes.

Definition 1 *Event E is* **common knowledge** *between agents 1 and 2 in state ω and at info-gap α if ω belongs to all the sets $K^n_{i,\alpha}(E)$ for $i = 1, 2$ and for all $n \geq 1$.*

We now discuss a fairly general theorem which shows that common knowledge is not possible at any positive info-gap.

The universe of discourse Ω is a normed space whose norm is denoted $\|\cdot\|$. A ball of radius r and centered at c is denoted $B_{r,c}$:

$$B_{r,c} = \{\omega : \ \|\omega - c\| \leq r\} \tag{33}$$

If $r < 0$ we define $B_{r,c} = \emptyset$.

Theorem 2 *Let $E \subseteq \Omega$ be contained in a ball $E^+ = B_{\rho,c}$. Two agents have info-gap models $\mathcal{U}_i(\alpha,\omega)$, $i = 1, 2$ for which there exist balls $B_{\rho_i(\alpha),\omega}$ such that:*

$$B_{\rho_i(\alpha),\omega} \subseteq \mathcal{U}_i(\alpha,\omega) \tag{34}$$

for all $\omega \in E^+$. The radius functions $\rho_i(\alpha)$ are both positive for $\alpha > 0$, and they do not depend upon ω. For $i = 3-j$ and for all $n \geq 1$, define the functions:

$$\sigma_{i,2n}(\alpha) = n[\rho_i(\alpha) + \rho_j(\alpha)] \tag{35}$$

$$\sigma_{i,2n-1}(\alpha) = n\rho_i(\alpha) + (n-1)\rho_j(\alpha) \tag{36}$$

With these conditions, the iterated knowledge functions of the two agents satisfy:

$$K^n_{i,\alpha}(E) \subseteq B_{\rho-\sigma_{i,n}(\alpha),c} \tag{37}$$

What this result means is that, if $\alpha > 0$, then there is an integer $n^*(\alpha)$ for which $K^n_{i,\alpha}(E) = \emptyset$ for all $n \geq n^*(\alpha)$. That is, a positive info-gap of size α precludes the possibility of common knowledge beyond the $n^*(\alpha)$th layer. Full common knowledge, to all levels of iteration n, is possible in the info-gap context only under full information: $\alpha = 0$. We can estimate $n^*(\alpha)$: it is no larger than the least value of n at which $\sigma_{i,n}(\alpha) > \rho$, where ρ is the radius of E^+, the ball containing E.

We see in theorem 2 a rather fundamental divergence between the classical theory of partitional information and the info-gap analog of that theory which is developed here. In the classical theory, an event E is common knowledge in state ω if ω belongs to all the iterated knowledge functions of E (Osborne and Rubinstein 1994, p.73). Such common knowledge is not only possible, but its existence is essential to such important results as Aumann's theorem (1976). In the info-gap case, however, complete common knowledge at all levels of iteration is not possible if $\alpha > 0$, since theorem 2 implies that no more than a finite number of knowledge functions are non-empty. Perhaps the ultimate source of the difference is the absence of the analog of property $P2$, eq.(2), from the info-gap theory, as explained in eqs.(4)–(8).

Example 3 Consider two agents $i = 1, 2$ whose info-gap models are identical balls:

$$\mathcal{U}_i(\alpha, \omega) = B_{\alpha,\omega}, \quad \alpha \geq 0 \tag{38}$$

Thus their information functions are the same.

Let $E = B_{\rho,c}$. The iterated knowledge function can be directly constructed, and is found to be:

$$K^n_{i,\alpha}(E) = \begin{cases} \emptyset & \alpha > \dfrac{\rho}{n} \\ B_{\rho-n\alpha,c} & \text{otherwise} \end{cases} \tag{39}$$

The set-inclusions in theorem 2 become equalities in this spherical example, and $n^*(\alpha) = \rho/\alpha$. While i's individual knowledge is identical to j's, common knowledge at info-gap α is possible only up to and including the nth layer for $n \leq \rho/\alpha$. ∎

Example 4 Prisoners' dilemma. A typical penalty function for the prisoners' dilemma (Davis, 1997) is shown in table 1 where 'C' means 'confess' and 'D' means 'don't confess'. For instance, in situation 'C, D' the prisoners receive penalties 1 and 3, respectively. Larger entries represent greater penalties. Each prisoner faces the dilemma that neither strategy, C or D, dominates the other. If one prisoner chooses C then the other would prefer C over D to obtain penalty 2 rather than 3. On the other hand, if one prisoner chooses D then the other would likewise prefer D to gain penalty 0 rather than 1. Since the prisoners can't collude, they each face a dilemma.

	C	D
C	2, 2	1, 3
D	3, 1	0, 0

Table 1: Typical penalty matrix for the prisoners' dilemma.

Now consider the situation in which neither player is certain about the penalties accruing under (C, C) and (D, D). That is, the penalty matrix is table 2. The values of x and y are uncertain to the prisoners, and can be any real numbers, positive and negative values being penalties and rewards, respectively.

	C	D
C	x, x	1, 3
D	3, 1	y, y

Table 2: Penalty matrix for the prisoners' dilemma with uncertainty.

Let Ω be $\Re^2$: the space of (x, y) values. Each prisoner's info-gap model for uncertainty in (x, y) is a ball, as in eq.(38).

Consider the event $E = B_{\rho,c}$ where c is the typical vector of penalty values: $c = (2, 0)$. Option D dominates C if and only if $x \geq 3$ and $y \leq 1$. Event E contains a penalty matrix in which C is dominated if and only if $\rho \geq 1$. The existence of D as a possibly dominating option might influence the prisoners' cogitations. Nominally (table 1) 'D, D' is the global optimum, though the choice of D is also the riskiest for each individual prisoner. If D dominates, however, it is Nash stable and both locally and globally optimal.

Recall from theorem 1 that one's knowledge decreases as α increases. Also, the nesting axiom of info-gap models states that the range of available contingencies increases as α increases. These observations motivate the interpretation that one's ampliative confidence decreases as α increases.

Let us consider $\rho \geq 1$, so D-dominance is possible (though not guaranteed) in E. We know from example 3 that the iterated knowledge function $K^n_{i,\alpha}(E)$ is not empty if and only if $n \leq \rho/\alpha$. If $\alpha > \rho$ then $K^1_{i,\alpha}(E) = \emptyset$, meaning that neither prisoner can infer, up to info-gap α, that E holds or that C is possibly dominated. If $\alpha \leq \rho$ then $K^1_{i,\alpha}(E) \neq \emptyset$ which means that each prisoner knows (with info-gap α) that a D-dominant penalty matrix is possible. However, if $\alpha > \rho/2$ then $K^2_{i,\alpha}(E) = \emptyset$, which means that neither prisoner can infer that the other prisoner can know that a D-dominant strategy is possible. Since collusion is the prisoners' best hope, this lack of common knowledge is significant. ∎

7 Common Knowledge, Coherent Uncertainties and Consensus

We have shown in theorem 2 that full common knowledge is not possible when accompanied by an info-gap. Consensus is nonetheless a possibility. We will show that consensus can occur if the info-gap models of the agents are coherent (in a sense to be defined). Furthermore we will show that coherence of the uncertainty models facilitates whatever common knowledge is possible. In summary, we establish, in an info-gap context, the connection between consensus and common knowledge. In this way we establish an info-gap analog of Aumann's famous result (1976).

$g(\alpha)$ and $h(\alpha)$ are **coherence functions** for two info-gap models, $\mathcal{U}_i(\alpha, \omega)$, $i = 1, 2$, if:

$$0 \leq g(\alpha) \leq h(\alpha) \tag{40}$$

and if:

$$\mathcal{U}_1(g(\alpha), \omega) \subseteq \mathcal{U}_2(\alpha, \omega) \subseteq \mathcal{U}_1(h(\alpha), \omega) \tag{41}$$

(Note the asymmetry of the indices 1 and 2.) $g(\alpha)$ and $h(\alpha)$ are, respectively, lower and upper coherence functions. If $g(\alpha) = h(\alpha)$ then $\mathcal{U}_1(g(\alpha), \omega) = \mathcal{U}_2(\alpha, \omega)$, which means that $\mathcal{U}_1$ is simply a scaled version of $\mathcal{U}_2$. If $g(\alpha)$ and $h(\alpha)$ are close then these info-gap models are similar though not identical in shape. Coherence functions are important in providing an indication of the possibility of agreement between the decision makers (Ben-Haim, 2001, chap. 9).

We defined the info-gap knowledge function $K_\alpha(E)$ in eq.(10), and the iterated info-gap knowledge functions $K^n_{i,\alpha}(E)$, in eq.(32). $K^n_{i,\alpha}(E)$ regards common knowledge: what i knows about what j knows about what i knows, etc. We now define an iterated self-knowledge function: what i knows about his own knowledge. For all $n \geq 1$ define:

$$K^n_{i,\alpha,\beta}(E) \quad = \quad K_{i,\alpha}(K^{n-1}_{i,\beta,\alpha}(E)) \tag{42}$$

where:

$$K^0_{i,\alpha,\beta}(E) = E \tag{43}$$

For instance, $K^2_{i,\alpha,\beta}(E)$ is the set of states in which i knows (up to info-gap α) the set of states in which i knows E (up to info-gap β).

Our first result establishes set-inclusion relations involving common knowledge, self-knowledge and coherence functions.

Lemma 4 *Given coherence functions $g(\alpha)$ and $h(\alpha)$ defined for info-gap models $\mathcal{U}_1(\alpha, \omega)$ and $\mathcal{U}_2(\alpha, \omega)$ in the sense of eq.(41), the knowledge functions satisfy:*

$$K^n_{1,h(\alpha),\alpha}(E) \;\subseteq\; K^n_{2,\alpha}(E) \;\subseteq\; K^n_{1,g(\alpha),\alpha}(E) \tag{44}$$

$$K^n_{1,\alpha,h(\alpha)}(E) \;\subseteq\; K^n_{1,\alpha}(E) \;\subseteq\; K^n_{1,\alpha,g(\alpha)}(E) \tag{45}$$

We have shown elsewhere (Ben-Haim, 2001, chap. 9) that, if $g(\alpha) = h(\alpha)$, then the two agents have identical robust-satisficing info-gap preferences on their options, regardless of the reward function. That is, uncertainty-coherence implies consensus. Furthermore we have shown that if $g(\alpha)$ and $h(\alpha)$ are 'close' then agreement is 'facilitated'. Relation (44) shows that:

$$g(\alpha) = h(\alpha) \quad \Longrightarrow \quad K^n_{2,\alpha}(E) = K^n_{1,g(\alpha),\alpha}(E) \text{ for all } n \tag{46}$$

That is, if the coherence functions are equal, then each common-knowledge function $K^n_{2,\alpha}(E)$ is equal to a self-knowledge function $K^n_{1,g(\alpha),\alpha}(E)$. If the info-gap models are completely coherent, then agent 1 can deduce all the accessible common knowledge from self knowledge alone. It is clear that if $g(\alpha) = h(\alpha)$, then the coherence functions defined in the other direction (2 and 1 interchanged in eq.(41)) are also equal. Lemma 4 now implies that agent 2 can also deduce all common knowledge from his own self-knowledge functions.

Furthermore, lemma 4 shows how 1's self-knowledge can be used to bracket his re-construction of common knowledge, when coherence between the info-gap models is less than complete.

To develop a result which is converse to implication (46) we first need lemma 5.

An info-gap model is **bounded** if all its sets $\mathcal{U}(\alpha, \omega)$ are bounded.

Lemma 5 *$\mathcal{U}_1(\alpha, \omega)$ and $\mathcal{U}_2(\alpha, \omega)$ are bounded info-gap models with knowledge functions $K_{1,\alpha}(E)$ and $K_{2,\alpha}(E)$, respectively.*

$$K_{2,\alpha}(E) = K_{1,\beta}(E) \quad \textit{for some } \omega_0 \in \Omega, \quad \text{for } E = \mathcal{U}_1(\beta, \omega_0) \text{ and for } \; E = \mathcal{U}_2(\alpha, \omega_0) \tag{47}$$

if and only if:

$$\mathcal{U}_2(\alpha, \omega) = \mathcal{U}_1(\beta, \omega) \text{ for all } \omega \tag{48}$$

This result says that the agents have the same knowledge at two particular events, though possibly at different levels of uncertainty, if and only if the info-gap uncertainty sets which comprise that knowledge are the same, at the respective levels of uncertainty. (Incidentally, using eq.(10), (48) implies that $K_{2,\alpha}(E) = K_{1,\beta}(E)$ for all E.)

Lemma 5 can be applied to show that:

$$K_{2,\alpha}(E) = K_{1,g(\alpha)}(E) \;\text{ for } E = \mathcal{U}_1(g(\alpha), \omega_0) \text{ and for } \; E = \mathcal{U}_2(\alpha, \omega_0) \tag{49}$$

if and only if:

$$\mathcal{U}_2(\alpha, \omega) = \mathcal{U}_1(g(\alpha), \omega) \text{ for all } \omega \tag{50}$$

If this holds for all $\alpha \geq 0$ then $g(\alpha)$ is both a lower and an upper coherence function. In other words, the equivalence of the knowledge functions of agents 1 and 2 in relation (49) implies complete coherence, $g(\alpha) = h(\alpha)$.

We collect the results of this discussion of lemmas 4 and 5 in the following theorem.

Theorem 3 *$\mathcal{U}_1(\alpha, \omega)$ and $\mathcal{U}_2(\alpha, \omega)$ are bounded info-gap models with knowledge functions $K_{1,\alpha}(E)$ and $K_{2,\alpha}(E)$, respectively, and with lower and upper coherence functions $g(\alpha)$ and $h(\alpha)$ in the sense of eq.(41). The following assertions are equivalent.*

$$g(\alpha) = h(\alpha) \quad \text{for all} \quad \alpha \tag{51}$$

$$K^n_{2,\alpha}(E) = K^n_{1,g(\alpha),\alpha}(E) \quad \text{for all } n,\ \alpha,\ E \tag{52}$$

For each value of α there is an ω_0 such that:

$$K_{2,\alpha}(E) = K_{1,g(\alpha)}(E) \quad \text{for} \quad E = \mathcal{U}_1(g(\alpha), \omega_0) \quad \text{and for} \quad E = \mathcal{U}_2(\alpha, \omega_0) \tag{53}$$

Aumann's theorem states that if two agents have the same prior probabilities and if their posterior probabilities are common knowledge, then these posteriors must be equal (Aumann, 1976; see also Osborne and Rubinstein, 1994, section 5.3). This equality must hold even if the agents' partitional information functions differ. Since their probability assessments are identical, the agents will presumably have a strong basis for further agreement on issues of judgment and action. In short, consensus can arise through common knowledge despite differing information.

A literal analog of Aumann's theorem is not possible in the info-gap context because there are no measure functions in an info-gap model of uncertainty, and because full common knowledge to all levels of iteration is not possible in the presence of an info-gap (theorem 2). Nonetheless, consensus is possible and theorem 3 establishes a connection between knowledge and agreement.

We know that if the info-gap models of two agents are coherent, that is, their lower and upper coherence functions are equal as in eq.(51), then their robust satisficing info-gap preferences will agree (Ben-Haim, 2001, section 9.1, corollary 2). Theorem 3 states that the info-gap models are coherent if and only if the knowledge functions are equal at particular events and at specific, perhaps differing, levels of info-gap uncertainty (eq.(53)). Lemma 5 shows that this equality of the knowledge functions is equivalent to identity of 'shape' though not 'scale' of the info-gap models. In other words, agents whose information differs only in level of uncertainty will agree on their preferences over the available actions (at any fixed level of satisficing reward).

We also know that if the coherence functions are 'close' (but not necessarily equal) at a point which depends on any pair of options, then the agents will agree on their preferences between these options (Ben-Haim, 2001, section 9.1, corollary 1). If the coherence functions differ, then so do the mutual and self-knowledge functions, as seen in eqs.(51) and (52) of theorem 3. But relation (44) of lemma 4 shows that if the coherence functions are close then these knowledge functions are also close. In short, consensus is possible if self-knowledge and common knowledge do not differ too greatly.

8 Contract Negotiation

In this section we apply theorem 3 to a negotiation between an employer and a prospective employee. For further discussion see (Ben-Haim, 2001, pp.240–243).

The problem. An employer wishes to offer a contract of employment to a prospective employee. The employee's activity will generate income η. The employee's wages $w(\eta, q)$ depend on the income according to some formula encoded in the decision vector q. The choice of q is the essence of the agreement which must be reached between employer and employee. In a bidding situation the employer hopes to make an offer, entailing a specification of q, which will be acceptable to the employee, and yet which will also assure adequate profit to the employer.

The income η depends on the effort ε which the employee will expend on the job, as well as on other unknown and uncontrollable factors. Nominally, the income is related to the employee's effort according to:

$$\eta = \gamma\varepsilon \tag{54}$$

where γ is a known quantity. Reality, of course, is more complicated. The employee declares an intended effort $\varepsilon_{\rm d}$. The employer may be uncertain about whether or not this declared intention reflects the true future effort to be exerted by the employee. Likewise, the employee may be overly optimistic (or deceptive) and hence uncertain as well about what effort will in fact be expended. Furthermore, the degree of disparity between effort and effectiveness is unclear to both parties. In short, the income is uncertain due to the uncertain effort and to additional unknown factors. The income and the effort are uncertain to both the employer and the employee, though their perceptions of these uncertainties are different.

The employer would like to know, before entering negotiations, or before actually specifying a class of wage formulas $w(\eta, q)$, that mutually acceptable agreement with an employee is plausible. A strong element of plausibility would be provided by coincidence of employee and employer preferences on the available q-values, despite their different perceptions of the uncertainties. This does not entirely guarantee that agreement will be reached, since the parties may disagree about the required minimal (satisficing) level of reward. That is, the preferences generated by an info-gap robustness function $\widehat{\alpha}(q, r_c)$ depend on the critical reward r_c, whose selection by the two sides may differ. Nonetheless, if the robustness functions for employer and employee generate identical preferences at any given value of r_c, then this is a firm basis for seeking a final agreement.

Uncertainty models. Let us suppose that the employer's uncertainty about the income η and employee effort ε is represented by the following info-gap model:

$$\mathcal{U}_2(\alpha, (\gamma\varepsilon_d, \varepsilon_d)) = \{(\eta, \varepsilon) : \quad \gamma\varepsilon(1-\alpha) \le \eta \le \gamma\varepsilon(1+\alpha) \\ \varepsilon_d(1-\alpha) \le \varepsilon \le \varepsilon_d(1+\alpha)\}, \quad \alpha \ge 0 \tag{55}$$

In other words, the employer's uncertainty about the income, η, varies as a symmetric interval of unknown size about the nominal value $\gamma\varepsilon$, where ε is the unknown effort exerted by the employee. In addition, the employer's uncertainty about the employee's effort ε varies in an unknown symmetric interval about the declared value ε_d.

The employee has a similar uncertainty model, with the single difference that the uncertain interval for the employee's effort is asymmetrical with respect to the same centerpoint. In particular, the interval is extended to lower values:

$$\mathcal{U}_1(\alpha, (\gamma\varepsilon_d, \varepsilon_d)) = \{(\eta, \varepsilon) : \quad \gamma\varepsilon(1-\alpha) \le \eta \le \gamma\varepsilon(1+\alpha) \\ \varepsilon_d(1-\psi\alpha) \le \varepsilon \le \varepsilon_d(1+\alpha)\}, \quad \alpha \ge 0 \tag{56}$$

where $\psi \ge 1$ is a value chosen by the employee to reflect an inclination to implement lower-than-declared values of effort ε.

Coherence functions and common knowledge. The maximal lower and minimal upper coherence functions, defined in the sense of eq.(41), are:

$$g(\alpha) = \frac{\alpha}{\psi} \quad \text{and} \quad h(\alpha) = \alpha \tag{57}$$

where $\psi \ge 1$ because the employee anticipates the possibility of expending less than the declared effort ε_d. The **incoherence,** $h(\alpha) - g(\alpha)$, of these info-gap models is:

$$\Delta = \frac{\psi - 1}{\psi}\alpha \tag{58}$$

which is non-negative and less than α.

We know from theorem 3 that, since the info-gap models are not coherent (that is, eq.(51) does not hold if $\psi > 1$), neither agent can precisely deduce the other's knowledge function from his own (eq.(53) also does not hold). Specifically, the disparity between $K_{1,g(\alpha)}(E)$ and $K_{2,\alpha}(E)$ increases with the incoherence, Δ.

As an example, let $E = \mathcal{U}_2(\alpha, \omega_0)$ for $\omega_0 = (\gamma\varepsilon_d, \varepsilon_d)$. The knowledge functions for this event are:

$$K_{2,\alpha}(E) = \{\omega_0\} \tag{59}$$

$$K_{1,g(\alpha)}(E) = \left\{(\gamma\varepsilon, \varepsilon) : \varepsilon \in \left[\varepsilon_d, \frac{1+\alpha}{1+\alpha-\Delta}\varepsilon_d\right]\right\} \tag{60}$$

These knowledge functions are equal if and only if the info-gap models are coherent ($\Delta = 0$). The upper bound of the interval of ε-values in (60) gets larger, and the knowledge functions become more different, as Δ increases towards α. As we expect from eq.(44) of lemma 4 for $n = 1$, $K_{2,\alpha}(E)$ is a subset of $K_{1,g(\alpha)}(E)$. However, if the employer (agent 2) tries to construct the employee's knowledge of E from 2's own self knowledge (which 2 could do if eq.(53) held), the employer errs increasingly as the incoherence rises.

Curiously, $K_{1,g(\alpha)}(E)$ grows as α, the info-gap, increases (recall from eq.(58) that Δ is proportional to α). This is surprising, in light of theorem 1 from which we know that knowledge is constricted as the info-gap grows as expressed by the set-inclusion in eq.(14). However, in the present example the event E also depends upon α, unlike theorem 1. So, in the present case $K_{1,g(\alpha)}(E)$ and $K_{2,\alpha}(E)$ become more different with increasing α: $K_{2,\alpha}(E)$, a singleton set, is invariant, while $K_{1,g(\alpha)}(E)$ grows. It becomes more difficult for 2 to estimate 1's knowledge as 1's knowledge grows.

9 Search and Evasion: Two Competitors

In a hunter's search after an evasive prey, the vector function $\omega(t)$ represents the displacement between hunter and prey as a function of time. Neither party knows $\omega(t)$, though each *could* know $\omega(t)$ if he knew his own as well as the other agent's strategy. In the absence of this knowledge, the evolution of $\omega(t)$ is shrouded in a cloud of possibilities whose range of variation increases with the agent's info-gap.

Agent 1 does not know 2's info-gap model for the uncertainty in $\omega(t)$. However, we will suppose that 1 has the resources to estimate coherence functions, in the sense of eq.(41), of the form:

$$g(\alpha) = \zeta\alpha, \quad h(\alpha) = \frac{1}{\zeta}\alpha, \quad 0 < \zeta \leq 1 \tag{61}$$

Agent 1 knows his own info-gap model for $\omega(t)$, so, knowing the value of ζ, 1 can bracket 2's knowledge functions based on eq.(44) of lemma 4. This is strategically important since it allows 1 to estimate what 2 knows, what 2 knows that 1 knows, etc. For instance, if 1 is going to try to force ω into a particular region E of Ω, 1 would like to know if 2 can detect this, and if 2 can know that 1 knows that ω is in E, etc. We will illustrate this with a simple example.

Agent 1 represents $\omega(t)$ as a truncated expansion of order M, the coefficients of which are uncertain real numbers but the expansion functions are known. Thus 1's info-gap model is a family of nested sets in $\Re^M$: the space of uncertain coefficient vectors. We consider an ellipsoid-bound info-gap uncertainty model in which the major axes of the ellipsoids are parallel to the coordinate axes:

$$\mathcal{U}_1(\alpha, \omega) = \left\{\nu : \ (\nu - \omega)^T D(\nu - \omega) \leq \alpha^2\right\}, \quad \alpha \geq 0 \tag{62}$$

where D is a diagonal real matrix with positive eigenvalues $\lambda_1, \ldots, \lambda_M$.

We will consider events which are rectangles in $\Re^M$. Define the rectangle centered at c and with side half-lengths s:

$$R(s, c) = \{\omega : \ |\omega_m - c_m| \leq s_m, \ m = 1, \ldots, M\} \tag{63}$$

which we define to be empty if $s_m < 0$ for any $m = 1, \ldots, M$.

The principal axes of the ellipsoid defining $\mathcal{U}_1(\alpha, \omega)$ are parallel to the sides of $R(s, c)$, and the the lengths of the semi-axes of this ellipsoid are $\alpha/\sqrt{\lambda_m}$, $m = 1, \ldots, M$. Thus $\mathcal{U}_1(\alpha, \omega)$ is contained within the rectangle $R(s, c)$ if ω is within the rectangle $R(s - \alpha\mu, c)$ where $\mu_m = 1/\sqrt{\lambda_m}$, $m = 1, \ldots, M$. That is, if $E = R(s, c)$, then 1's knowledge function for E is:

$$K_{1,\alpha}(E) = R(s - \alpha\mu, c) \tag{64}$$

which is empty if $s_m < \alpha/\sqrt{\lambda_m}$ for any m. Eq.(64) states that the knowledge function of a rectangle is a somewhat smaller rectangle. This enables the easy derivation of the self-knowledge functions $K^n_{1,\alpha,\beta}(E)$, as we now show.

Define the M-vector function whose mth element $\delta_m(n, \alpha, \beta)$ is:

$$\delta_m(2n, \alpha, \beta) = \frac{n(\alpha + \beta)}{\sqrt{\lambda_m}} \tag{65}$$

$$\delta_m(2n + 1, \alpha, \beta) = \frac{(n + 1)\alpha + n\beta}{\sqrt{\lambda_m}} \tag{66}$$

By induction one can show that, for $E = R(s, c)$, 1's self-knowledge functions are:

$$K^n_{1,\alpha,\beta}(E) = R[s - \delta(n, \alpha, \beta),\ c] \tag{67}$$

which is empty if and only if $s_m < \delta_m(n, \alpha, \beta)$ for some $m = 1, \ldots, M$.

Define:

$$\rho = \min_{1 \le m \le M} s_m \sqrt{\lambda_m} \tag{68}$$

which is the least ratio of the side half-length of the rectangle $R(s, c)$ to the corresponding semi-axis of the unit ellipsoid. Recalling $g(\alpha)$ in eq.(61) and using eq.(67) we see that, for $E = R(s, c)$:

$$K^{2n}_{1,g(\alpha),\alpha}(E) \quad = \quad \emptyset \quad \text{if and only if} \quad n > \frac{\rho}{\alpha(\zeta + 1)} \tag{69}$$

$$K^{2n+1}_{1,g(\alpha),\alpha}(E) \quad = \quad \emptyset \quad \text{if and only if} \quad n > \frac{\rho - \zeta\alpha}{\alpha(\zeta + 1)} \tag{70}$$

For instance, for $n = 1$ in eq.(69), we see from this and eq.(44) that $K^2_{2,\alpha}(E) = \emptyset$, meaning that 2 cannot know that 1 knows E at info-gap α, if and only if $1 > \frac{\rho}{\alpha(\zeta+1)}$. However $K_{1,\alpha}(E) = R(s - \alpha\mu,\ c)$ which is *not* empty, meaning that 1 *can* know E at info-gap α, if and only if $\alpha \le \rho$. These two conditions can occur simultaneously:

$$1 + \zeta > \frac{\rho}{\alpha} \ge 1 \tag{71}$$

which is strategically significant for 1. When (71) holds, agent 1 can know E at info-gap α, but 2 cannot know (at the same info-gap) that 1 knows E at info-gap α. From 1's point of view, (71) defines a "beneficial" level of uncertainty, α, which 1 can use against 2.

It is not surprising that the left inequality in (71) can fail if ζ is too small, since a small ζ means that 1's info-gap model is highly incoherent with 2's model.

It is interesting to note that the left inequality in (71) can also fail if the info-gap is too small, while the right inequality fails if α is too large. Both observations are explained by theorem 1: knowledge constricts as α grows. When α is too large 1 cannot know E at all, while when α is small enough 2 can know that 1 knows E.

10 Teamwork and the Need to Know

In section 9 we showed how one agent can exploit uncertainty to his advantage against an opponent by monitoring the access to common knowledge. We now consider an inversion of the search-and-evasion problem; we study a situation in which coordinated effort between teams requires the transmission of information. Since such transfer is costly, we explore the questions of how much and what sort of information must be transmitted. We formulate this in the context of an example.

Consider two design teams. T_2, the electronics team, is responsible for developing an electronics card. T_1, the thermal-hydraulics team, must design the system for cooling the card.

In order to complete its task, T_1 must know the 3 linear dimensions of the card and its total power consumption. Let $\omega \in \Omega$ denote the vector of these 4 quantities. In fact, T_1 can proceed with only approximate knowledge of ω. This knowledge can be approximate in two senses: knowledge of a set $E \subset \Omega$ rather than knowledge of ω itself, and knowledge of the set E with info-gap uncertainty $\alpha > 0$

rather than with absolute certainty. T_1 has an info-gap model for its uncertainty in ω: $\mathcal{U}_1(\alpha, \omega)$, $\alpha \geq 0$.

The electronics folks, T_2, also do not know ω, as well as many other aspects of the card design. In fact, the domain of discourse in which T_2 operates is much more complicated than Ω. Furthermore, T_2's domain of discourse may not be separable into the form $\Omega \times \Psi$, where Ψ is the space of the additional variables which concern T_2. This means that T_2's info-gap model need *not* be a Cartesian product $\mathcal{U}_{2,\omega}(\alpha, \omega) \times \mathcal{U}_{2,\psi}(\alpha, \psi)$ of uncertainty about ω with uncertainty about all the other factors ψ which concern T_2. If this separability were to hold, T_2 could simply hand the model $\mathcal{U}_{2,\omega}(\alpha, \omega)$ over to T_1. Even if this is possible, however, $\mathcal{U}_{2,\omega}(\alpha, \omega)$ may be far more complex than T_1 needs in order to make its design decisions. $\mathcal{U}_{2,\omega}(\alpha, \omega)$ may be some elaborate and highly informative info-gap model, while T_1 only needs a simple interval or ellipsoidal model.

In liaison between the teams it is sufficient, as far as T_1 is concerned, to talk only in terms of the 4-vector ω, and only in terms of an info-gap model whose level of sophistication meets T_1's design needs. This raises the following possibility: T_2 will periodically give to T_1 coherence functions, in the sense of eq.(40), which correspond roughly but adequately (for T_1's purposes) to T_2's improving knowledge. That is, these coherence functions represent the fidelity of T_1's info-gap model to a simplified version of T_2's model.

This scheme for knowledge transfer will be satisfactory if the following conditions can be satisfied. $E \subset \Omega$ is a particular event which, when it happens:

• T_1 must be able to infer that E holds, up to uncertainty α. This is because E is a 'trigger' event of design-significance to T_1.

• T_2 must be able to infer that T_1 is able to infer that E holds, up to uncertainty α. This is for purposes of monitoring and control. T_2 must be able to know that T_1 is now triggered to act.

The info-gap, α, plays a central role. The knowledge we are dealing with is anticipation or estimation of unknown and possibly future reality. For instance, T_2 believes, with uncertainty α, that E holds, where E specifies a class of final designs.

To see how this works let us adopt the formalism of section 9. T_1's info-gap model for uncertainty in ω is eq.(62). T_2 periodically provides T_1 with coherence functions of the form of eq.(61). Consider events E in Ω which are rectangles $R(s, c)$ as in eq.(63). T_1's self-knowledge functions $K^n_{1,\alpha,\beta}(E)$ are also rectangles as in eq.(67).

T_1 can infer E, in the presence of uncertainty α, if $K_{1,\alpha}(E) \neq \emptyset$. $K_{1,\alpha}(E) = R(s - \alpha\mu, c)$, eq.(64), so this holds if $s_m \geq \alpha/\sqrt{\lambda_m}$ for all $m = 1, \ldots, M$. Defining ρ as in eq.(68), this condition is $\alpha \leq \rho$.

T_2 can infer (up to info-gap α) that T_1 can infer E (up to α) if $K^2_{2,\alpha}(E) \neq \emptyset$. To avoid using T_2's complicated info-gap models, we employ the lefthand inclusion in eq.(44) of lemma 4. $K^2_{2,\alpha}(E)$ is not empty if $K^2_{1,h(\alpha),\alpha}(E) \neq \emptyset$, where $h(\alpha)$ is the upper coherence function, $h(\alpha) = \alpha/\zeta$. We find that $K^2_{1,h(\alpha),\alpha}(E) = R[s - \delta(2, h(\alpha), \alpha), c]$ is not empty if:

$$\alpha \leq \rho \frac{\zeta}{1 + \zeta} \tag{72}$$

which is stricter than the condition on $K_{1,\alpha}(E)$. This relation is a constraint on the ambient uncertainty under which adequate knowledge transfer occurs when T_2 supplies T_1 with coherence functions alone.

A large value of ρ facilitates the success of the knowledge transfer, since then the conditions on the two teams' knowledge can be satisfied even in the presence of large info-gap. Recall that ρ is the least ratio of the side half-length of the rectanglar trigger-event $E = R(s, c)$ to the corresponding semi-axis of the unit ellipsoid of T_1's info-gap model. ρ is large if T_1's uncertainty tends to be small compared to the trigger-event E.

A large value of ζ, and hence of $\zeta/(1 + \zeta)$, likewise facilitates successful knowledge transfer. ζ varies from 0 to 1 as T_1's info-gap model increases in coherence with T_2's uncertainty about ω. Here we face a trade-off. On the one hand, T_1 needs only a simplied model as compared to T_2's model, so coherence will tend to be low. On the other hand, this limits the uncertainty-range within which transfer of coherence functions is adequate.

11 Appendix: Proofs

Proofs for section 4

Proof of lemma 1. By the contraction axiom:

$$0 \in \mathcal{U}(0,0) \tag{73}$$

By the translation axiom:

$$\mathcal{U}(0,\omega) = \mathcal{U}(0,0) + \omega \tag{74}$$

From which, with the nesting axiom, we have:

$$\omega \in \mathcal{U}(0,\omega) \subseteq \mathcal{U}(\alpha,\omega) \tag{75}$$

which completes the proof. ∎

Proof of lemma 2. $K1$ results from the fact that $\mathcal{U}(\alpha,\omega) \subseteq \Omega$.

$K2$:

$$\omega \in K_\alpha(E) \ \text{ implies that } \ \mathcal{U}(\alpha,\omega) \subseteq E \tag{76}$$

Hence:

$$E \subseteq F \ \text{ implies that } \ \mathcal{U}(\alpha,\omega) \subseteq F \tag{77}$$

Thus $\omega \in K_\alpha(F)$.

$K3$: By the definition of $K_\alpha(E)$ in eq.(10) we have:

$$\begin{aligned} K_\alpha(E) \cap K_\alpha(F) &= \{\omega : \ \mathcal{U}(\alpha,\omega) \subseteq E \text{ and } \mathcal{U}(\alpha,\omega) \subseteq F\} && (78)\\ &= \{\omega : \ \mathcal{U}(\alpha,\omega) \subseteq E \cap F\} && (79)\\ &= K_\alpha(E \cap F) && (80)\end{aligned}$$

which completes the proof that $K_\alpha(E)$ obeys the analog of $K3$. ∎

Proof of lemma 3. For any $\omega \in K_\alpha(E)$, the definition of $K_\alpha(E)$ and the contraction and translation axioms of info-gap models imply that:

$$\omega \in \mathcal{U}(\alpha,\omega) \ \text{ and } \ \mathcal{U}(\alpha,\omega) \subseteq E \tag{81}$$

Hence $\omega \in E$ which completes the proof. ∎

Proofs for section 6

Proof of theorem 2. We first need a few elementary results which we present without proof.

$$\begin{aligned} B \subseteq B^+ \quad &\Longrightarrow \quad \{\omega : \ A(\omega) \subseteq B\} \subseteq \{\omega : \ A(\omega) \subseteq B^+\} && (82)\\ A^-(\omega) \subseteq A(\omega) \quad &\Longrightarrow \quad \{\omega : \ A(\omega) \subseteq B\} \subseteq \{\omega : \ A^-(\omega) \subseteq B\} && (83)\end{aligned}$$

An elementary property of balls is, for $r \geq 0$:

$$\{\omega : \ B_{r,\omega} \subseteq B_{\rho,c}\} = B_{\rho-r,c} \tag{84}$$

Our proof of eq.(37) will be an induction on n.

We first prove eq.(37) for $n = 1$. By definition:

$$K_{i,\alpha}(E) = \{\omega : \mathcal{U}_i(\alpha, \omega) \subseteq E\} \tag{85}$$

From eqs.(34) and (83):

$$\{\omega : \mathcal{U}_i(\alpha, \omega) \subseteq E\} \subseteq \left\{\omega : B_{\rho_i(\alpha),\omega} \subseteq E\right\} \tag{86}$$

From relation (82):

$$\left\{\omega : B_{\rho_i(\alpha),\omega} \subseteq E\right\} \subseteq \left\{\omega : B_{\rho_i(\alpha),\omega} \subseteq E^+\right\} \tag{87}$$

Hence:

$$K_{i,\alpha}(E) \subseteq \left\{\omega : B_{\rho_i(\alpha),\omega} \subseteq E^+\right\} \tag{88}$$

With eq.(84), and recalling that $E^+ = B_{\rho,c}$:

$$K_{i,\alpha}(E) \subseteq B_{\rho-\rho_i(\alpha),c} \tag{89}$$

which proves eq.(37) for $n = 1$.

Now suppose that eq.(37) holds for n. For $n + 1$:

$$\begin{aligned} K^{n+1}_{i,\alpha}(E) &= K_{i,\alpha}(K^n_{j,\alpha}(E)) && (90)\\ &= \left\{\omega : \mathcal{U}_i(\alpha, \omega) \subseteq K^n_{j,\alpha}(E)\right\} && (91) \end{aligned}$$

By eqs.(82) and (83) and the inductive hypothesis:

$$K^{n+1}_{i,\alpha}(E) \subseteq \left\{\omega : B_{\rho_i(\alpha),\omega} \subseteq B_{\rho-\sigma_{j,n}(\alpha),c}\right\} \tag{92}$$

Thus by eq.(84):

$$K^{n+1}_{i,\alpha}(E) \subseteq B_{\rho-\sigma_{j,n}(\alpha)-\rho_i(\alpha),c} \tag{93}$$

If $n = 2k$:

$$\sigma_{j,n}(\alpha) + \rho_i(\alpha) = k\rho_j(\alpha) + k\rho_i(\alpha) + \rho_i(\alpha) = \sigma_{i,2(k+1)-1}(\alpha) = \sigma_{i,n+1}(\alpha) \tag{94}$$

If $n = 2k - 1$:

$$\sigma_{j,n}(\alpha) + \rho_i(\alpha) = k\rho_j(\alpha) + (k-1)\rho_i(\alpha) + \rho_i(\alpha) = \sigma_{i,2k}(\alpha) = \sigma_{i,n+1}(\alpha) \tag{95}$$

So eq.(37) is confirmed inductively. ■

Proofs for section 7

Proof of lemma 4. Our proof is by induction on n. We will rely extensively upon eqs.(82) and (83) and theorem 1.

Eq.(44) for $n = 1$. By definition:

$$\begin{aligned} K^1_{2,\alpha}(E) = K_{2,\alpha}(E) &= \{\omega : \mathcal{U}_2(\alpha, \omega) \subseteq E\} && (96)\\ K^1_{1,h(\alpha),\alpha}(E) = K_{1,h(\alpha)}(E) &= \{\omega : \mathcal{U}_1(h(\alpha), \omega) \subseteq E\} && (97)\\ K^1_{1,g(\alpha),\alpha}(E) = K_{1,g(\alpha)}(E) &= \{\omega : \mathcal{U}_1(g(\alpha), \omega) \subseteq E\} && (98) \end{aligned}$$

Employing the definition of the coherence functions $g(\alpha)$ and $h(\alpha)$ in eq.(40), the nesting of info-gap models, and relation (83) we see from eqs.(96)–(98) that:

$$\begin{aligned} K^1_{1,h(\alpha),\alpha}(E) &\subseteq K^1_{2,\alpha}(E) && (99)\\ K^1_{2,\alpha}(E) &\subseteq K^1_{1,g(\alpha),\alpha}(E) && (100) \end{aligned}$$

which proves eq.(44) for $n = 1$.

Eq.(45) for $n = 1$: each of the three sets is, by definition, equal to $K_{1,\alpha}(E)$. Thus lemma 4 is proven for $n = 1$.

Now suppose the lemma holds for some $n \geq 1$. We first prove the righthand side of relation (45) for $n + 1$. From (44) for n:

$$K^n_{2,\alpha}(E) \subseteq K^n_{1,g(\alpha),\alpha,}(E) \tag{101}$$

This, together with property $K2$ of lemma 2, shows that:

$$K_{1,\alpha}(K^n_{2,\alpha}(E)) \subseteq K_{1,\alpha}(K^n_{1,g(\alpha),\alpha,}(E)) \tag{102}$$

This, by definition, is:

$$K^{n+1}_{1,\alpha}(E) \subseteq K^{n+1}_{1,\alpha,g(\alpha)}(E) \tag{103}$$

which is the righthand inclusion in (45) at $n + 1$.

We now prove the lefthand side of relation (45) for $n + 1$. From (44) for n:

$$K^n_{1,h(\alpha),\alpha,}(E) \subseteq K^n_{2,\alpha}(E) \tag{104}$$

This, together with property $K2$ of lemma 2, shows that:

$$K_{1,\alpha}(K^n_{1,h(\alpha),\alpha,}(E)) \subseteq K_{1,\alpha}(K^n_{2,\alpha}(E)) \tag{105}$$

This, by definition, is:

$$K^{n+1}_{1,\alpha,h(\alpha)}(E) \subseteq K^{n+1}_{1,\alpha}(E)) \tag{106}$$

which is the lefthand inclusion in (45) at $n + 1$ which completes the proof of (45).

We now prove first the righthand and then the lefthand side of relation (44) for $n + 1$. From (44) for $n = 1$ for any set F:

$$K_{2,\alpha}(F) \subseteq K_{1,g(\alpha)}(F) \tag{107}$$

Choose $F = K^n_{1,\alpha,g(\alpha)}(E)$ so this becomes:

$$K_{2,\alpha}(K^n_{1,\alpha,g(\alpha)}(E)) \subseteq K_{1,g(\alpha)}(K^n_{1,\alpha,g(\alpha)}(E)) \tag{108}$$

By (45) at n we have:

$$K^n_{1,\alpha}(E) \subseteq K^n_{1,\alpha,g(\alpha)}(E) \tag{109}$$

Combining this with property $K2$ of lemma 2, shows that:

$$K_{2,\alpha}(K^n_{1,\alpha}(E)) \subseteq K_{2,\alpha}(K^n_{1,\alpha,g(\alpha)}(E)) \tag{110}$$

Thus (108) becomes:

$$K_{2,\alpha}(K^n_{1,\alpha}(E)) \subseteq K_{1,g(\alpha)}(K^n_{1,\alpha,g(\alpha)}(E)) \tag{111}$$

This, by definition, is:

$$K^{n+1}_{2,\alpha}(E) \subseteq K^{n+1}_{1,g(\alpha),\alpha}(E)) \tag{112}$$

which is the righthand inclusion in (44) at $n + 1$.

We now prove the lefthand side of relation (44) for $n + 1$. From (44) for $n = 1$ for any set F:

$$K_{1,h(\alpha)}(F) \subseteq K_{2,\alpha}(F) \tag{113}$$

Choose $F = K^n_{1,\alpha,h(\alpha)}(E)$ so this becomes:

$$K_{1,h(\alpha)}(K^n_{1,\alpha,h(\alpha)}(E)) \subseteq K_{2,\alpha}(K^n_{1,\alpha,h(\alpha)}(E)) \tag{114}$$

By (45) at n we have:

$$K^n_{1,\alpha,h(\alpha)}(E) \subseteq K^n_{1,\alpha}(E) \quad (115)$$

Combining this with property $K2$ of lemma 2, shows that:

$$K_{2,\alpha}(K^n_{1,\alpha,h(\alpha)}(E)) \subseteq K_{2,\alpha}(K^n_{1,\alpha}(E)) \quad (116)$$

Thus (114) becomes:

$$K_{1,h(\alpha)}(K^n_{1,\alpha,h(\alpha)}(E)) \subseteq K_{2,\alpha}(K^n_{1,\alpha}(E)) \quad (117)$$

This, by definition, is:

$$K^{n+1}_{1,h(\alpha),\alpha}(E) \subseteq K^{n+1}_{2,\alpha}(E) \quad (118)$$

which is the righthand inclusion in (44) at $n+1$. This completes the proof of lemma 4. ∎

In order to prove lemma 5 we first need the following result.

Lemma 6 *$\mathcal{U}(\alpha,\omega)$ is a bounded info-gap model.*

$$\mathcal{U}(\alpha,\omega) \subseteq \mathcal{U}(\alpha,\omega_1) \quad \Longrightarrow \quad \omega = \omega_1 \quad (119)$$

Proof of lemma 6. Subtracting ω from both sides of the inclusion in the statement of the lemma, and using the translation axiom, yields:

$$\mathcal{U}(\alpha,0) \subseteq \mathcal{U}(\alpha,\omega_1-\omega) \quad (120)$$

Likewise, subtracting ω_1 rather than ω yields:

$$\mathcal{U}(\alpha,\omega-\omega_1) \subseteq \mathcal{U}(\alpha,0) \quad (121)$$

Hence:

$$\mathcal{U}(\alpha,\omega-\omega_1) \subseteq \mathcal{U}(\alpha,0) \subseteq \mathcal{U}(\alpha,\omega_1-\omega) \quad (122)$$

Adopt the inductive hypothesis:

$$\mathcal{U}(\alpha,2^n(\omega-\omega_1)) \subseteq \mathcal{U}(\alpha,0) \subseteq \mathcal{U}(\alpha,2^n(\omega_1-\omega)) \quad (123)$$

which we have proven true for $n=0$. Subtract $2^n(\omega-\omega_1)$ from the two outer sets in (123), resulting in:

$$\mathcal{U}(\alpha,0) \subseteq \mathcal{U}(\alpha,2^{n+1}(\omega_1-\omega)) \quad (124)$$

Likewise, subtract $2^n(\omega_1-\omega)$ from the two outer sets in (123), resulting in:

$$\mathcal{U}(\alpha,2^{n+1}(\omega-\omega_1)) \subseteq \mathcal{U}(\alpha,0) \quad (125)$$

Inclusions (124) and (125) prove the inductive hypothesis, (123), for all $n \ge 0$.

Now, since $\mathcal{U}(\alpha,0)$ is bounded there is a number ρ such that $\mathcal{U}(\alpha,\omega)$ is contained in a ball of radius ρ centered at 0:

$$\mathcal{U}(\alpha,0) \subseteq B_{\rho,0} \quad (126)$$

The lefthand inclusion in (123) implies that:

$$2^n(\omega-\omega_1) \in B_{\rho,0} \quad (127)$$

for all $n \ge 0$. Hence $\omega = \omega_1$. ∎

Proof of lemma 5. That eq.(48) implies (47) derives from eq.(10) by substitution.

Now consider the reverse implication. Choose $E = \mathcal{U}_1(\beta, \omega_0)$. Employing the definition of $K_{1,\beta}(E)$ and lemma 6 we find:

$$K_{1,\beta}(E) = \{\omega : \mathcal{U}_1(\beta, \omega) \subseteq \mathcal{U}_1(\beta, \omega_0)\} = \{\omega_0\} \tag{128}$$

By the supposition of the lemma:

$$K_{2,\alpha}(E) = K_{1,\beta}(E) = \{\omega_0\} \tag{129}$$

which, by the definition of $K_{2,\alpha}(E)$, implies that:

$$\mathcal{U}_2(\alpha, \omega_0) \subseteq \mathcal{U}_1(\beta, \omega_0) \tag{130}$$

By a similar argument, interchanging α with β, 1 with 2, and using $E = \mathcal{U}(\alpha, \omega_0)$, we can prove that:

$$\mathcal{U}_1(\beta, \omega_0) \subseteq \mathcal{U}_2(\alpha, \omega_0) \tag{131}$$

The last two relations, together with the translation axiom, show that:

$$\mathcal{U}_1(\beta, \omega) = \mathcal{U}_2(\alpha, \omega) \tag{132}$$

for arbitrary ω. This completes the proof. ■

Proof of theorem 3. Eq.(51): Eq.(51) implies eq.(52) by lemma 4 and eq.(46). Eq.(52) implies Eq.(53) by inclusion.

Eq.(53): Eq.(53) implies eq.(51) by lemma 5 and eqs.(49) and (50). Eq.(51) implies eq.(52).

Eq.(52): Eq.(52) implies eq.(53) by inclusion, and eq.(53) implies. eq.(51). ■

12 References

Robert J. Aumann, 1976, Agreeing to disagree, *Annals of Statistics*, 4:1236–1239. Reprinted in Robert J. Aumann, *Collected Papers,* vol. 1, MIT Press, 2000, pp.593–596.

Yakov Ben-Haim, 1999, Set-models of information-gap uncertainty: Axioms and an inference scheme, *Journal of the Franklin Institute,* 336: 1093–1117.

Yakov Ben-Haim, 2001, *Information-gap Decision Theory: Decisions Under Severe Uncertainty,* Academic Press. Appearing in the *Academic Press Book Series on Decision and Risk.*

Morton D. Davis, 1997, *Game Theory: A Nontechnical Introduction,* Dover Publications, reprint of 1983 Basic Books edition.

Stuart Berg Flexner, 1980, *The Random House Dictionary,* New York.

Drew Fudenberg, and Jean Tirole, 1991, *Game Theory,* MIT Press.

Frank H. Knight, 1921, *Risk, Uncertainty and Profit.* Houghton Mifflin Co. Re-issued by University of Chicago Press, 1971.

Henry E. Kyburg, jr., 1990, Getting fancy with probability, *Synthese,* 90: 189–203.

Dov Monderer and Dov Samet, 1989, Approximating common knowledge with common beliefs, *Games and Economic Behavior,* 1: 170–190.

Martin J. Osborne and Ariel Rubinstein, 1994, *A Course in Game Theory,* MIT Press.

Oxford English Dictionary, 1999, 2nd edition on CD-ROM, Version 2.0. 'Knowledge', *n.,* II.5.a.

G.L.S. Shackle, 1972, *Epistemics and Economics: A Critique of Economic Doctrines,* Cambridge University Press, re-issued by Transaction Publishers, 1992.