Common Knowledge
and State-Dependent Equilibria

Nuh Aygun Dalkiran[1], Moshe Hoffman[2], Ramamohan Paturi[3],
Daniel Ricketts[3], and Andrea Vattani[3]

[1] Bilkent University, Department of Economics
dalkiran@bilkent.edu.tr
[2] The Rady School of Management & Department of Computer Science and
Engineering, UC San Diego
mohoffman@ucsd.edu
[3] Department of Computer Science and Engineering, UC San Diego
{paturi,daricket,avattani}@ucsd.edu

Abstract. Many puzzling social behaviors, such as avoiding eye contact, using innuendos, and insignificant events that trigger revolutions, seem to relate to common knowledge and coordination, but the exact relationship has yet to be formalized. Herein, we present such a formalization. We state necessary and sufficient conditions for what we call state-dependent equilibria – equilibria where players play different strategies in different states of the world. In particular, if everybody behaves a certain way (e.g. does not revolt) in the usual state of the world, then in order for players to be able to behave a different way (e.g. revolt) in another state of the world, it is both necessary and sufficient for it to be common p-believed that it is not the usual state of the world, where common p-belief is a relaxation of common knowledge introduced by Monderer and Samet [16]. Our framework applies to many player r-coordination games – a generalization of coordination games that we introduce – and common (r, p)-beliefs – a generalization of common p-beliefs that we introduce. We then apply these theorems to two particular signaling structures to obtain novel results.

1 Introduction

In the popular parable "The Emperor's New Clothes" [2], a gathering of adults pretends to be impressed by the Emperor's dazzling new suit despite the fact that he is actually naked. It is not until an innocent child cries out "But he has nothing on at all!" that the Emperor's position of authority and respect is questioned. This is a metaphor for a number of common political situations in which the populace knows the current regime is inept but takes no action against it until some seemingly insignificant event occurs, such as the child's cry. In fact, in Tunisia, despite years of political repression and poverty, it was not until the previously unknown street vendor Mohamed Bouazizi set himself on fire that citizens rose up in protest. Common knowledge – everyone knows

that everyone knows that... – might offer such an explanation for this strange phenomenon: while the boy's cry and the self immolation of Mohamed Bouazizi do not teach anyone that the government is inept, they make it commonly known that the government is inept. Likewise, common knowledge has been proffered as an explanation for many other puzzling social behaviors: it is common to avoid eye-contact when caught in an inappropriate act, despite the fact that looking away, if anything, increases the conspicuousness of a shameful deed. Nevertheless, even Capuchin monkeys look away when they ignore a request to help an ally in a tussle [21]. And few adults after a nice date are fooled by the inquiry "Would you like to come upstairs for a drink?" yet all but the most audacious avoid the explicit request [22].

Many authors have aptly noted that common knowledge plays an important role in these puzzling social behaviors [10, 4, 6, 22]. Avoiding eye contact prevents common knowledge that you were noticed, using innuendos enables a speaker to request something inappropriate without making the request commonly known, and prohibiting public displays of criticism of the government while not preventing people from realizing the flaws of their government, prevent the flaws from being commonly known. Authors have argued that common knowledge is important in these situations because common knowledge is needed for coordination. But without formal arguments, many important questions still remain, such as: what exactly needs to be "commonly known" in order to "coordinate"? What exactly will happen in the absence of common knowledge? Miscoordination? When common knowledge is lacking, but almost present, e.g. if everyone is pretty sure that everyone is pretty sure... will this have the same effect as common knowledge? Such details, which may seem pedantic, are crucial for answering practical questions such as: if I cannot think up an innuendo, will an appropriately placed cough midsentence do the trick? Why is it that sometimes we use innuendos and sometimes we go out of our way to state the obvious?

We will formalize the role of common knowledge in coordination, which will enable us to address each of these questions. The crucial step in our formalism is based on the insight of Rubinstein [23]. Rubinstein considers coordination games – games in which players make choices such that they would like to mimic the choice that others make. Rubinstein supposes that players coordinate on a particular action A in a given situation. He then supposes that the situation changes and asks whether the players can coordinate on a different action instead. He shows that unless it is commonly known that the situation has changed, players still must coordinate on A. The intuition is clear: even if one player knows that circumstances have changed, if he thinks the other player does not know this, then he expects the other player to play as if circumstances have not changed. Since it is a coordination game, he best responds by playing as if circumstances have not changed. Likewise, even if both players know that circumstances have changed, and both players know that both players know this, but one player does not realize the second player has this second degree of knowledge, then this player will expect the other player to play as if circumstances have not changed.

By the above argument, he best responds by playing as if circumstances have not changed. The same logic continues indefinitely.

Rubinstein presents a particular instance in which the above logic holds. The contribution of our paper is to show that this logic holds quite generally, for any two player coordination game, and in fact, for a generalization to many players. And moreover, we show that common knowledge is not just necessary for changing behaviors when circumstances change, but common knowledge is also sufficient. We hope that this will lead to a deeper understanding of these puzzling social behaviors, as well as some novel predictions.

OUR RESULTS. In this paper, we introduce *state-dependent equilibria*, which we define as equilibrium strategies in which players take different actions when the circumstances change. This notion allows us to address the questions that were left unanswered by the informal discussions of common knowledge and coordination. In particular, we characterize the conditions under which rational players are able to play state-dependent equilibria.

We begin by considering two-player coordination games. We show that it is not quite common knowledge that determines the existence of state-dependent equilibria but rather a relaxation of common knowledge. This notion corresponds with common p-beliefs, as developed by Monderer and Samet [16]: each believes with probability at least p that each believe with probability at least p.... In our framework, we show that p depends on the precise payoffs of the game and corresponds to the risk dominance of Harsanyi and Selten [14].

We then introduce a natural n-player generalization of coordination games that we call *r-coordination games* in which coordination on an action is successful if at least some fraction r of the players take that action. Accordingly, we also develop a generalization of common p-beliefs for this setting.

In order to derive our results, we provide a unifying theoretical framework for analyzing our games. Our framework gives tight necessary and sufficient conditions on the players' beliefs under which a state-dependent equilibrium exists. These conditions depend on the payoffs of the game (in particular on the risk dominance) and, in the case of r-coordination games, on the threshold fraction r required for successful coordination.

Our final contribution is to apply this framework both to simple but puzzling social behavior and to more complex distributed phenomena that arise in biology, economics, and sociology. The first application is eye-contact. We offer a post hoc explanation for why we avoid eye-contact when caught in an inappropriate act. For the second and third applications, we show how our results can be applied to situations in which the true state of the world is observed by all players with arbitrarily small noise, as in the global games literature [18, 20, 19]. This yields some novel predictions about social behaviors, such as which cues can be used to instigate a revolution, and when a researcher's reputation can be resilient to substandard work.

Due to space constraints, the proofs of our claims will appear in the full version of the paper.

1.1   Related Work

The concept of common knowledge was first formalized in multi-modal logic in 1969 by Lewis [15]. Aumann later put common knowledge in a set-theoretic framework [3].

In 1989, Rubinstein used common knowledge to analyze a problem related to the coordinated attack problem in computer science [23]. This problem, called the Electronic Mail Game, was the first example that common knowledge is very different than any finite order of knowledge. Rubinstein showed that the lack of common knowledge prevents players from switching strategies (i.e. prevents the existence of state-dependent equilibria) in the Electronic Mail Game. See [17] for a retrospective on the Electronic Mail Game. Our results show that common knowledge is not just necessary but also sufficient and holds for any coordination game and even r-coordination games.

Carlsson and Van Damme showed that when players have noisy signals about the payoffs in a coordination game, as the noise vanishes, the unique equilibrium in the game becomes the risk dominant equilibrium [5]. Morris and Shin applied this result to bank runs and currency crisis, showing that there is a unique underlying value at which currencies collapse and bank runs occur, in contrast to previous models, which permitted multiple equilibria and prevented comparative static analysis [18, 20, 19]. In some of our applications, we use similar signaling structures, but the uncertainty does not affect the payoffs. We find circumstances under which no state-dependent equilibria exist.

Monderer and Samet developed an approximate notion of common knowledge called common p-beliefs, which is relevant in our framework. We will draw heavily on their definitions and results [16]. Others have discussed the role of common knowledge in social puzzles, albeit less formally than in the aforementioned literature. Chwe discusses the role of common knowledge in public rituals [6]. Pinker *et al* discusses the role of common knowledge in innuendos [22]. Binmore and Friedell discuss the role of common knowledge in eye contact [9, 4]. In our paper we formalize the role of common knowledge in many of these social puzzles.

The role of common knowledge has been studied in the fields of distributed computing and artificial intelligence [11, 7, 12]. This line of work suggests that knowledge is an important abstraction for distributed systems and for the design and analysis of distributed protocols, in particular for achieving consistent simultaneous actions. Fagin and Halpern [13, 8] present an abstract model for knowledge and probability in which they assign to each agent-state pair a probability space to be used when computing the probability that a formula is true. A complexity-theoretic version of Aumann's celebrated Agreement Theorem is provided in [1].

2   Preliminaries

We will adopt the set-theoretic formulation of common knowledge introduced by Aumann [3]. In this model, there is a set Ω of "states of the world". Each player i has some information regarding the true state of the world. This information is

given by a partition Π_i of Ω. In particular, for $\omega \in \Omega$, $\Pi_i(\omega)$ is the set of states indistinguishable from ω to player i – that is, when ω occurs, player i knows that one of the states in $\Pi_i(\omega)$ occurred but not which one. Finally, there is a probability distribution μ over Ω, representing the (common) prior belief of the players over the states of the world. These parameters all together constitute the information structure.

Definition 1 (Information structure). *An information structure is a tuple $\mathcal{I} = (N, \Omega, \mu, \{\Pi_i\}_{i \in N})$ where N is the set of players (with $n := |N|$), Ω is the set of possible states of the world, μ is a strictly positive common prior probability distribution over Ω, and Π_i is the information partition of player i. $\Pi_i(\omega)$ gives the set of states indistinguishable from ω to player i.*

A (Bayesian) game is now defined by an information structure, a set of possible actions for each player and a state-dependent utility for each player.

Definition 2 (Bayesian game). *A Bayesian game Γ is a tuple $(\mathcal{I}, \{A_i\}_{i \in N}, \{u_i\}_{i \in N})$ where $\mathcal{I} = (N, \Omega, \mu, \{\Pi_i\}_{i \in N})$ is an information structure, A_i is the (finite) set of possible actions that player i can take, $u_i : A_1 \times \ldots \times A_n \times \Omega \to \mathcal{R}$ is the utility for player i given the state of the world and the actions of all players.*

A strategy profile prescribes the action (possibly randomized) that each player takes at each state of the world.

Definition 3 (Strategy profile). *A strategy profile is a function $\sigma = (\sigma_1, \ldots, \sigma_n) : \Omega \to A_1 \times \ldots \times A_n$ that specifies what action each player takes in each state of the world.*

Since a player cannot distinguish between states belonging to the same partition, it is enforced that if a player i plays some strategy $\sigma = \sigma_i(\omega)$ at some state $\omega \in \Omega$, it must be the case that i plays σ at all states $\omega' \in \Pi_i(\omega)$. We can now recall the definition of Bayesian Nash equilibrium.

Definition 4 (Bayesian Nash equilibrium). *A strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n) : \Omega \to A_1 \times \ldots \times A_n$ is a Bayesian Nash equilibrium (BNE) of Γ if for all $i \in N$,*

1. *$\sigma_i(\omega) = \sigma_i(\omega')$ whenever $\omega \in \Pi_i(\omega')$.*
2. *$\int_{\omega \in \Omega} u_i(\sigma_i(\omega), \sigma_{-i}(\omega)) \mathrm{d}\mu(\omega) \geq \int_{\omega \in \Omega} u_i(\sigma_i'(\omega), \sigma_{-i}(\omega)) \mathrm{d}\mu(\omega)$ for all σ' satisfying property 1.*

We now introduce our key definition of *state-dependent equilibria*, which we define as equilibrium strategies in which players take different actions when the circumstances change. This notion allows us to address the questions that were left unanswered by the informal discussions of common knowledge and coordination.

Definition 5 (State-dependent BNE). *We say that a Bayesian Nash equilibrium σ^* is state-dependent if for some $\omega, \omega' \in \Omega$, $i \in N$, we have that $\sigma_i^*(\omega) = A$ and $\sigma_i^*(\omega') = B$.*

We now define the notion of p-belief, introduced by Monderer and Samet [16], which extends the notion of common knowledge by Aumann [3]. Let p be a number between 0 and 1. We say that a player i p-believes the event E at state of the world ω if the subjective probability that i assigns to E at ω is at least p. That is, whenever ω is the true state of the world, i believes that an event in E occurred with probability at least p. Henceforth, we will use short expressions such as "i p-believes E at ω" to refer to this concept.

We denote by $\mathcal{B}_i^p(E)$ the set of all states of the world at which player i p-believes E.

Definition 6 (p-belief [16]). *For any $0 \leq p \leq 1$, we say that player i p-believes E at ω if $\mu(E \mid \Pi_i(\omega)) \geq p$. We will denote by $\mathcal{B}_i^p(E)$ the event that i p-believes E, i.e. $\mathcal{B}_i^p(E) = \{\omega \mid \mu(E \mid \Pi_i(\omega)) \geq p\}$.*

Observe that by definition of $\mathcal{B}_i^p(E)$, the notation $\omega \in \mathcal{B}_i^p(E)$ indicates that whenever ω occurs, player i believes with probability at least p that the event E occurred. An event E is then defined p-evident if whenever it occurs, each player i believes with probability at least p that it indeed occurred.

Definition 7 (evident p-belief [16]). *An event E is evident p-belief if for all $i \in N$ we have $E \subseteq \mathcal{B}_i^p(E)$.*

The following concept extends the notion of common knowledge.

Definition 8 (common p-belief [16]). *An event C is common p-belief at state ω if there exists an evident p-belief event E such that $\omega \in E$, and for all $i \in N$, $E \subseteq \mathcal{B}_i^p(C)$.*

Monderer and Samet provide a nice example that illustrates this concept: suppose the true state is either E or F with equal probability. The true state is announced and each of two players independently hears the announcement with probability $1 - \epsilon$, $0 < \epsilon < 1/2$. Then if E is the true state and both hear the announcement then E is common p-belief for all $p < 1 - \epsilon$ even though it is not common knowledge.

3   Two Player Framework

In this section we consider the classic 2-player, 2-strategy symmetric coordination game. The payoffs are as follows:

	A	B
A	a, a	b, c
B	c, b	d, d

Assumption 1 (Coordination game). *We make the following standard assumption on the parameters of a symmetric coordination game: $a > c$ and $d > b$.*

Throughout this paper, we will use $p^* = \frac{d-b}{d-b+a-c}$. This value is called risk-dominance [14]. Note that if player i believes with probability exactly p^* that the other player will play A at ω, then player i will be indifferent between playing A and B at ω.

For convenience, we will use the following definitions throughout this section.

Definition 9. *Given any strategy profile σ, we let $A_i(\sigma) = \{\omega | \sigma_i(\omega) = A\}$ and $B_i(\sigma) = \{\omega | \sigma_i(\omega) = B\}$, i.e. the set of states where player i plays A and B respectively.*

We now state our main result for the 2-player case. The main question we ask is when is it possible for the two players to coordinate on different actions in different states of the world. We answer this question in terms of the existence of evident p-belief events (where p depends on the payoff matrix) showing that such events are necessary and sufficient.

Theorem 1. *There exists a state-dependent Bayesian Nash equilibrium σ^* if and only if there exists a non-empty evident p^*-belief event E and a non-empty evident $(1 - p^*)$-belief event F such that $E \cap F = \emptyset$.*

While evident knowledge is both necessary and sufficient for state-dependent equilibria, our theorem further allows us to specify how the strategies must depend on these evident events, which we express in the following corollary:

Corollary 1. *A strategy profile σ^* is a state-dependent Bayesian Nash equilibrium if and only if there exists a non-empty evident p^*-belief event E and a non-empty evident $(1 - p^*)$-belief event F such that $\mathcal{B}_i^{p^*}(E) \cap \mathcal{B}_i^{1-p^*}(F) = \emptyset$ and $\mathcal{B}_i^{p^*}(E) \cup \mathcal{B}_i^{1-p^*}(F) = \Omega$ for all i, in which case $A_i(\sigma^*) = \mathcal{B}_i^{p^*}(E)$ and $B_i(\sigma^*) = \mathcal{B}_i^{1-p^*}(F)$ for all i.*

Our next corollary states the relationship between state-dependent equilibria and common knowledge.

Corollary 2. *If σ^* is a Bayesian Nash equilibrium such that $\sigma_i^*(\omega) = A$ and $\sigma_i^*(\omega') = B$, then $\neg \omega'$ is common p^*-belief at ω and $\neg \omega$ is common $(1-p^*)$-belief at ω'.*

4   Application: A Rationale for Avoiding Eye-Contact

Two Charedi men, Michael and Dave, go to a bar, and each spots the other, purposely looking away before meeting eyes. Why?

Suppose that the next day they have to decide whether to tell the Rabbi. If one expects the other to tell, he is better off also admitting to his actions. On the other hand, if one does not expect the other to tell, then he is better off also not admitting to his transgression. The payoffs can be interpreted as the coordination game from the two-player framework by interpreting A as the act of not telling the Rabbi, B as the act of telling the Rabbi.

We make the reasonable assumption that if at least one of the men stays home, neither tells the Rabbi that he saw the other player at the bar (since he in fact did not). We will use our framework from section 3 to show that (a) there is always an equilibrium in which they both tell the Rabbi if they make eye-contact at the bar, and (b) under mild assumptions, if they do not make eye contact, neither will tell the Rabbi.

THE MODEL. We now specify the information structure: we suppose that in one state of the world, at least one of them stays home ($\mathcal{H}$) while in another state of the world, Dave enters the bar, and Michael is already sitting at the bar. When Dave walks in, Michael is either staring at the bartender, in which case he would not see Dave, or looking at the door, in which case he would. As soon as Dave enters, he sees Michael, so he quickly turns around and walks out. Dave turns around before or after noticing if Michael saw him.

The set of possible states of the world is given by $\Omega = \{\mathcal{H}, (\mathcal{M}, \mathcal{D}), (\mathcal{M}', \mathcal{D}), (\mathcal{M}, \mathcal{D}'), (\mathcal{M}', \mathcal{D}')\}$. We interpret the states of the world as follows: $\mathcal{H}$ is the state where Dave *does not* go to the bar and stays at ($\mathcal{H}$)ome. $\mathcal{M}$ is the event that Michael goes to the bar and sees Dave, and $\mathcal{D}$ is the event that Dave sees Michael. $(\mathcal{M}, \mathcal{D})$ is the state where Dave goes to the bar, Michael sees him, and Dave sees that Michael saw him (i.e. they make eye-contact). $(\mathcal{M}', \mathcal{D}')$ is the state where Dave goes to the bar, Michael is looking at the bartender, and Dave leaves the bar before checking if Michael saw him.

The information partitions are given as follows:

$$\Pi_{\mathcal{M}} = \{\{\mathcal{H}, (\mathcal{M}', \mathcal{D}'), (\mathcal{M}', \mathcal{D})\}, \{(\mathcal{M}, \mathcal{D}')\}, \{(\mathcal{M}, \mathcal{D})\}\}$$
$$\Pi_{\mathcal{D}} = \{\{\mathcal{H}\}, \{(\mathcal{M}', \mathcal{D}'), (\mathcal{M}, \mathcal{D}')\}, \{(\mathcal{M}', \mathcal{D})\}, \{(\mathcal{M}, \mathcal{D})\}\}$$

Observe that $(\mathcal{M}, \mathcal{D})$ is an evident p^*-belief event, that is, when eye contact happens, it becomes common knowledge between Michael and Dave as expected.

We use the following independent probabilities to deduce the priors over the state space: p_B is the probability that Dave goes to the bar i.e., he *does not* stay home; $p_{\mathcal{M}'}$ is the probability that Michael is looking at the bartender when Dave walks in; $p_{\mathcal{D}'}$ is the probability that, conditioned on Dave going to the bar, he leaves the bar without noticing Michael.

Our first claim is an almost trivial one which shows that there always exists an equilibrium in which they both tell the Rabbi if they make eye-contact.

Claim. There exists a Bayesian-Nash equilibrium of Γ such that $\sigma^*(\mathcal{H}) = (A, A)$ and $\sigma^*((\mathcal{M}, \mathcal{D})) = (B, B)$ for any $p_B, p_{\mathcal{M}'}, p_{\mathcal{D}'}$.

Our next claim shows conditions under which if Michael and Dave do not make eye-contact, they must continue playing A if they play A on $\mathcal{H}$. That is, suppose Michael and Dave coordinate on (A, A) when Dave stays home; under what conditions is it the case that they can play (B, B) only at $(\mathcal{M}, \mathcal{D})$, i.e. only when they make eye-contact.

Claim. Suppose σ^* is a Bayesian-Nash equilibrium of Γ with $\sigma^*(\mathcal{H}) = (A, A)$. If $p_{\mathcal{M}'} > p^*$ and $\frac{p_B p_{\mathcal{M}'}}{p_B p_{\mathcal{M}'} + (1 - p_B)} < 1 - p^*$ then $\sigma^*(\omega) \neq (B, B)$ for all $\omega \neq (\mathcal{M}, \mathcal{D})$.

Now that we have formalized why someone might want to avoid eye-contact, we can discuss when this is worthwhile. For instance, avoiding eye contact will not serve any purpose when it is very likely that they saw each other, e.g. if the bar had nobody else present and was very well lit (i.e. when $p_{\mathcal{M}'}$ and $p_{\mathcal{D}'}$ are small). Likewise, avoiding eye contact serves no purpose if, when it is commonly known that both parties see each other doing an act, neither is expected to play any differently than

if neither transgressed (i.e. $\sigma^*(\mathcal{M}, \mathcal{D}) = (A, A)$). For example, the transgression is not perceived as related to the ensuing coordination game, e.g. if the two religious men have already discussed their secret abhorrence of the religion.

Moreover, avoiding eye contact only serves a purpose if there will be an ensuing coordination game (i.e. $a > c$). If in fact Michael would prefer to rat on Dave, regardless of whether Dave rats on Michael (e.g. because he knows the Rabbi will believe him, and he would like Dave to be excommunicated) then Dave does not help himself by avoiding Michael's eyes. In fact, to the extent that Dave thinks this might be the case, he might want to avoid eye contact, as it may make his presence more conspicuous to Michael.

Lastly, Michael may even purposely make eye contact, or yell out "hey Dave, is that you," if he in fact *wants* to switch from them both playing A to both playing B (which would be the case if $d > a$). For instance, this would be the case if Dave was looking for someone to leave the community with him and help him start a new life in the secular world.

5 n-Player Framework

We now introduce r-coordination games. Let Ω be all possible states of the world. There are n players, each of whom can take action A or B. A player's payoff for a particular action is a function of the fraction of players who play B. In particular, a player's payoffs are a function of whether the fraction of players who play B exceeds a threshold $\bar{r}$. Let r denote the fraction of players who play B. The payoffs are as follows.

$$u_i(A, r) = \begin{cases} a : r \leq \bar{r} \\ b : r > \bar{r} \end{cases} \qquad u_i(B, r) = \begin{cases} c : r \leq \bar{r} \\ d : r > \bar{r} \end{cases}$$

We again use assumption 1 on the values of the parameters, namely that $a > c$ and $d > b$. In this context, these assumptions on the payoff parameters generalize that of a 2 player coordination game in that a player best respond by playing A if and only if sufficiently many others play A.

We will also assume that n is sufficiently large such that a particular player's decision to play A or B does not affect whether r exceeds $\bar{r}$.

Furthermore, we will again use $p^* = \frac{d-b}{d-b+a-c}$. For n-players, p^* is a generalization of risk dominance. If player i believes with probability exactly p that at least $(1 - \bar{r})$ players will play A at ω, then player i will be indifferent between playing A and B at ω.

Note that this setup is a generalization of the two player setup. In particular, if there are two players, then we can let $\bar{r}$ be any value in $(1/2, 1)$ in order to obtain the two player model.

In Definitions 10, 11, and 12, we generalize p-beliefs, evident p-beliefs, and common p-beliefs to n players.

Definition 10 ((r, p)-belief). *For any $0 \leq p \leq 1$ and any $0 \leq r \leq 1$, we say that event E is (r, p)-belief at ω if $|\{i \mid \omega \in \mathcal{B}_i^p(E)\}| \geq rn$. We define $\mathcal{B}^{r,p}(E) = \{\omega : |\{i \mid \omega \in \mathcal{B}_i^p(E)\}| \geq rn\}$ as the event that at least a fraction of r players p-believes E.*

Definition 11 (evident (r, p)-belief). *An event E is evident (r, p)-belief if $E \subseteq \mathcal{B}^{r,p}(E)$.*

Definition 12 (common (r, p)-belief). *Given an event C, let $C^0 = \mathcal{B}^{r,p}(C)$ and inductively define $C^n = \mathcal{B}^{r,p}(\bigcap_{i<n} C^i)$ for all $n \geq 2$. Then C is common (r, p)-belief at ω if $\omega \in \bigcap_{n \geq 1} C^n$*

Note that common (r, p)-beliefs is identical to common p-beliefs when $n = 2$ and $r = 1$. The following theorem and corollaries are analogous to our two-player theorems and corollaries, despite the differing setup and proofs.

Theorem 2. *There exists a state-dependent Bayesian Nash equilibrium σ^* if and only if there exists a non-empty evident $(1 - \bar{r}, p^*)$-belief event E and a non-empty evident $(s, 1 - p^*)$-belief event F such that $E \cap F = \emptyset$ for some $s > \bar{r}$.*

Corollary 3. *A strategy profile σ^* is a state-dependent Bayesian Nash equilibrium if and only if there exists a non-empty evident $(1 - \bar{r}, p^*)$-belief event E and a non-empty evident $(s, 1 - p^*)$-belief event F for some $s > \bar{r}$ such that $\mathcal{B}_i^{p^*}(E) \cap \mathcal{B}_i^{1-p^*}(F) = \emptyset$ and $\mathcal{B}_i^{p^*}(E) \cup \mathcal{B}_i^{1-p^*}(F) = \Omega$ for all i, in which case $A_i(\sigma^*) = \mathcal{B}_i^{p^*}(E)$ and $B_i(\sigma^*) = \mathcal{B}_i^{1-p^*}(F)$ for all i.*

Corollary 4. *If σ^* is a Bayesian Nash equilibrium such that $|\{j \mid \sigma_j^*(\omega) = A\}| \geq 1 - \bar{r}$ and $|\{j \mid \sigma_j^*(\omega') = B\}| > \bar{r}$, then $\neg \omega'$ is common $(1 - \bar{r}, p^*)$-belief at ω and $\neg \omega$ is common $(\bar{r}, 1 - p^*)$-belief at ω'.*

6 n-Player Application: The Emperor's Clothes

Suppose that John Doe is on his way to being the next game theorist superstar. He finally comes out with his first paper, and superficially it is a spectacular paper. However, the paper offers no real insight, a fact that John attempts to hide with mathematical complexity. And this is fairly clear to nearly everyone in the field. Nevertheless, editors start requesting the paper, departments start offering him positions, conferences start asking him to give the keynote. Why?

Presumably, no one wants to be the lone person in the field who disrespects the superstar. For example, nobody wants to be the only person *not* to invite John to a conference or a special journal issue; he might end up with a powerful enemy, even if John's research is not good. However, if everyone in the field disrespects John Doe, then everyone benefits from doing likewise, since no one wants his keynote speaker to be unpopular or his new recruit never to be invited to conferences. Thus, we can model this as a r-coordination game where A is the act of showing John Doe respect (e.g. inviting him to a conference), and B is an act of disrespect.[1]

We make the assumption that if in fact John Doe's research were as great as people expected, then everyone would treat him with respect. Furthermore, we assume

---

[1] Note that "The Emperor's New Clothes" can be seen as a metaphor for this story. John Doe is analogous to the Emperor and his colleagues are analogous to the citizens who do not, initially, publicly disrespect the obviously flawed superstar.

that if a person can detect that John's research is bad, he can only approximately estimate how many others can detect this as well. We will show (Theorem 3) that, under mild conditions, if John's research is bad, no matter what fraction of people in the field can detect that his research has no insight, he will still be treated with respect. However, if people know exactly what fraction of the field know that John's research is bad, and that fraction is sufficiently high, then it is possible to treat John with disrespect (subsequent Claim). This is in stark contrast with the case where the error in a person's estimate is arbitrarily small.

THE MODEL. We model the information structure as follows: we assume that if John's research is in fact bad, then $1 - \epsilon$ of the population can detect that it is bad. Everyone who can detect that it is bad has some impression of how easy it is for others to detect how bad it is; namely, they each get a signal θ_i which is independently drawn from $\mathcal{U}[\epsilon - \delta, \epsilon + \delta]$. After observing his private signal, but not ϵ, player i can choose to play A or B. As in the general setup, the payoff from each action is a function of the fraction of players who play B. Let r denote the fraction of players who play B. The payoffs are as in Section 5.

We can interpret Theorem 3 as follows. Suppose players disrespect John if their private signal θ_i of the true state ϵ is smaller than some (arbitrarily small) threshold $\bar{\epsilon}$. Then, if the fraction $\bar{r}$ of players needed to coordinate on B is larger than the risk-dominance p^*, this set of strategies is not an equilibrium. Note that the condition on $\bar{r}$ does *not* depend on $\bar{\epsilon}$. Another way of interpreting our results is the following. Even if many believe that many believe that many believe...that John's research is bad (for finitely many iterations), John will still be respected. Whereas, if it is common knowledge (subsequent Claim), e.g. if it is publicly announced how bad John's research is, he will no longer be respected.

Theorem 3. *Let $\epsilon \sim \mathcal{U}[0,1]$ and $\theta_i \sim^{iid} \mathcal{U}[\epsilon - \delta, \epsilon + \delta]$ for all i and for some $\delta > 0$. Let σ^* be a strategy profile such that $\sigma_i^*(\theta_i) = B$ when $\theta_i \leq \bar{\epsilon}$ and $\sigma_i^*(\theta_i) = A$ when $\theta_i > \bar{\epsilon}$ for some $\bar{\epsilon} \in [\delta, 1 - \delta]$. Then for $\delta \to 0$, σ^* is not a Bayesian Nash equilibrium if $\bar{r} > p^*$.*

We contrast this result with the scenario in which the exact value of ϵ is observed by those who can detect that John's research is bad (i.e. $\theta_i = \epsilon$). The following claim can be easily established.

Claim. The strategy profile σ^* is a Bayesian Nash equilibrium if $\sigma_i^*(\theta_i) = A$ if $\epsilon \leq 1 - \bar{r}$ and $\sigma_i^*(\theta_i) = B$ otherwise.

References

[1] Aaronson, S.: The complexity of agreement. In: Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing, STOC 2005, pp. 634–643. ACM, New York (2005)

[2] Andersen, H.C.: The Emperor's New Clothes. Fairy Tales Told for Children - First Collection (1837)

[3] Aumann, R.J.: Agreeing to Disagree. The Annals of Statistics 4(6), 1236–1239 (1976)

[4] Binmore, K.: Game theory - a very short introduction. Clarendon Press (2007)

[5] Carlsson, H., van Damme, E.: Global Games and Equilibrium Selection. Econometrica 61(5), 989–1018 (1993)

[6] Chwe, M.S.-Y.: Rational Ritual: Culture, Coordination, and Common Knowledge. Princeton University Press (2003)

[7] Dwork, C., Moses, Y.: Knowledge and common knowledge in a byzantine environment: Crash failures. Information and Computation 88(2), 156–186 (1990)

[8] Fagin, R., Halpern, J.Y.: Reasoning about knowledge and probability. J. ACM 41(2), 340–367 (1994)

[9] Friedell, M.F.: On the structure of shared awareness. Behavioral Science 14(1), 28–39 (1969)

[10] Geanakoplos, J.: Common knowledge. Journal of Economic Perspectives 6(4), 53–82 (1992)

[11] Halpern, J.Y., Moses, Y.: Knowledge and common knowledge in a distributed environment. J. ACM 37(3), 549–587 (1990)

[12] Halpern, J.Y., Moses, Y.: A guide to completeness and complexity for modal logics of knowledge and belief. Artif. Intell. 54(3), 319–379 (1992)

[13] Halpern, J.Y., Tuttle, M.R.: Knowledge, probability, and adversaries. J. ACM 40(4), 917–960 (1993)

[14] Harsanyi, J.C., Selten, R.: A General Theory of Equilibrium Selection in Games. MIT Press Books, vol. 1. The MIT Press (1988)

[15] Lewis, D.: Convention: A Philosophical Study. Harvard University Press, Cambridge (1969)

[16] Monderer, D., Samet, D.: Approximating common knowledge with common beliefs. Games and Economic Behavior 1(2), 170–190 (1989)

[17] Morris, S.: Coordination, communication and common knowledge: A retrospective on electronic mail game. Oxford Review of Economic Policy 18(4) (February 2002)

[18] Morris, S., Shin, H.S., Yale University. Cowles Foundation for Research in Economics: Global games: theory and applications. Cowles Foundation discussion paper. Cowles Foundation for Research in Economics (2000)

[19] Morris, S., Shin, H.S.: Unique equilibrium in a model of self-fulfilling currency attacks. American Economic Review 88(3), 587–597 (1998)

[20] Morris, S., Shin, H.S.: Rethinking multiple equilibria in macroeconomic modeling. In: NBER Macroeconomics Annual 2000, vol. 15, NBER Chapters, pp. 139–182. National Bureau of Economic Research, Inc. (June 2001)

[21] Perry, S., Manson, J.: Manipulative Monkeys: The Capuchins of Lomas Barbudal. Harvard University Press (2008)

[22] Pinker, S., Nowak, M.A., Lee, J.J.: The logic of indirect speech. PNAS 105(3), 833–838 (2008)

[23] Rubinstein, A.: The electronic mail game: Strategic behavior under "almost common knowledge". American Economic Review 79(3), 385–391 (1989)