C O P E

Member since 2018
JM13714

Richmond H. Thomason*

COMMON KNOWLEDGE, COMMON ATTITUDES, AND SOCIAL REASONING

Abstract

For as long as there have been theories about common knowledge, they have been exposed to a certain amount of skepticism. Recent more sophisticated arguments question whether agents can acquire common attitudes and whether they are needed in social reasoning. I argue that this skepticism arises from assumptions about practical reasoning that, considered in themselves, are at worst implausible and at best controversial. A proper approach to the acquisition of attitudes and their deployment in decision making leaves room for common attitudes. Postulating them is no worse off than similar idealizations that are usefully made in logic and economics.

Keywords: Common knowledge, belief, nonmonotonic logic, practical reasoning.

1. Preliminaries

Many approaches to coordinated practical reasoning rely on common knowledge—reciprocal and iterative knowledge of a proposition by a group.[1]

---

*The author wishes to thank two anonymous referees for helpful comments on this paper.

[1] *Warning:* Here and in the title to this paper, I speak of "knowledge" and "common knowledge." This is because that is the familiar term, and it will be easier for readers to identify the topic of this paper if I begin with what is familiar. But what I really have in mind—and this is crucially important—is common belief, or better, common supposition for some particular practical purpose.

These include game theory, bargaining theory, the theory of communications protocols, the theory of distributed computing, the theory of multiagent systems, the analysis of convention, the theory of grounding in human and machine communication, and various specific applications.[2]

All of these applications have to do with practical reasoning in group situations. The issue of whether common knowledge and similar attitudes are legitimate in social reasoning is a special case of the question of how propositional attitudes figure in practical reasoning of any sort. I believe that it can't be properly understood without situating it in this more general arena.

Despite its acceptance in many theoretical circles, second thoughts about common knowledge come readily to mind. We can understand what it means to ask whether someone knows that we know that they know something, but only with a certain amount of difficulty. And, with further iterations, the difficulty rapidly increases beyond anyone's intuitive capability. Perhaps this is why [20][p. 246] says "as most commentators would agree, mutual knowledge* is from the point of view of psychological reality at best problematic."[3]

Schiffer doesn't say what commentators he has in mind, and may be thinking of personal communications. The only published examples I'm aware of prior to 2017 are [22, 23, 24], which criticize the appropriateness of common knowledge in accounting for conversational common ground. [7] asks how children could acquire the concept. This is a legitimate and perhaps challenging question, but is hardly a criticism.

[23][p. 18] raises perhaps the first objection that would occur to a critic: to establish common knowledge, speakers "would have, in principle, to perform an infinite series of checks." If the point is that a generalization involving infinitely many instances can't be concluded without considering

---

[2]For game and bargaining theory, see [1, 6]. For distributed computing, see [16]. For multiagent systems, see [13, 2]. For convention, see [11]. For the theory of communication, see [3, 25]. For an application, see [15].

[3]Authors have been slow to coordinate on a terminology; some use "common" others use "mutual." Schiffer adds an asterisk to indicate, apparently, that he finds the notion artificial, although his formalization doesn't deviate in important respects from others. More recently, some economists have added to the confusion by using "mutual knowledge" for knowledge simply shared by a group. According to this terminology, common knowledge would be the limit of all finite iterations of mutual knowledge. In this paper, I myself will use "common" for this limiting notion. I'll avoid the term "mutual," except in quoting authors who use it.

each instance separately, it is clearly wrong: we can know that for each number there is a larger prime number without having to think about each case. Construed this way, the objection does raise the technical challenge of showing that conclusions about common knowledge can be derived from a finite axiomatic basis: [11, 19], for instance, address this issue.

If the point is that it is psychologically difficult to think about even moderately complex iterations of the knowledge operator, the prime number theorem also provides a counterexample, because it is quite difficult to show that large numbers are prime.

But there remains a more problematic version of this objection. The proof of the prime number theorem invokes a uniform method, applicable to any number. But with common knowledge, each finite iteration is defeasible and can depend on new and different evidence. Take the example of a simple public announcement: Ann says to Bob, "I'll be at home this afternoon." To know after the announcement that Bob knows she'll be at home, Ann needs to assume Bob heard and understood her. For Bob to know that Ann knows that Bob knows she'll be at home, Bob needs to assume that Ann knows that he heard and understood what she said; and so forth. In principle, this series of knowledge claims could hold up to any n but fail at $n+1$; this circumstance is particularly salient in the *coordinated attack scenario*, which we'll consider in Section 6.1. This version of the objection raises a technical challenge that has not been adequately addressed in the literature. I'll argue in Section 4 that using a nonmonotonic epistemic logic will solve the problem.

2. Lederman's challenge

Sperber and Wilson raise another, slightly different objection: that "the assumption of mutual knowledge may always be mistaken," [24][p. 19]. The objection amounts to this: in many situations calling for mutual knowledge, the conditions for knowledge simply don't exist. This objection is skeptical, and calls for a philosophical remedy rather than a technical response. But it will be instructive to frame the response in connection with more philosophically sophisticated versions of the objection, presented by Harvey Lederman in two recent articles,[4]

---

[4] [9, 10]. And [8] provides a useful survey of the relevant issues and literature.

Lederman concentrates on practical decision-making rather than conversational common ground, and questions the value of common attitudes in accounting for publicity and in modeling many cases of interpersonal reasoning. His arguments purport to show that it is impossible in practice for a rational agent to acquire such attitudes. If this were so, then it certainly would make no sense for interpersonal deliberation to require common attitudes. Moreover, Lederman points to experimental evidence that seems to show, in some cases at least, that humans arrive at decisions without obtaining common beliefs, much less common knowledge.

3.   Common attitudes and practical reasoning

Before addressing Lederman's arguments in detail, we need to consider the much more general issue of the relation between propositional attitudes and decision making. These strategic considerations matter: some models of belief acquisition make it hard to see how common beliefs could be acquired. And some views about reasons for action make it hard to see how mutual beliefs can serve in this capacity.

Many authors—especially computer scientists—who work with common attitudes speak somewhat recklessly in terms of "knowledge" when, if they were more philosophically minded, they would use "belief." Both knowledge and belief figure in decision making—but knowledge is more appropriate for evaluating decisions once they have been made. An agent who is criticizing another's or her own earlier decision is in a position to separate knowledge from belief. But for a deliberating agent, it is practical conviction that supports action, and it doesn't matter for the decision-making process whether this conviction actually counts as knowledge when the decision is evaluated.

So far, this should not be very controversial. More controversial, perhaps, is the idea that practical acceptance is not a matter of bringing a general-purpose, previously prepared attitude to bear, but is tailored *ad hoc* to the decision at hand, and depends not only on purely evidential factors but on risk.

As far as I know, the idea that agents tailor their practical attitudes to the specific decision-making context first appears in [21]. There, Simon proposed that the standards for an acceptable solution can depend may be adjusted during the deliberative process. In [26, 30], I claim this is also

true of practical beliefs; thresholds for activating beliefs in a deliberative context are adjusted according to estimates of the risk of acting on the belief.

The most familiar examples of this phenomenon are cases of *subtractive risk sensitivity*, where a belief disappears upon the realization that it would be risky to act on it. For instance, I normally believe that people receive email messages I send, and because nothing is unusual about a message I sent to my accountant, I believe she received it. But when I realize I may miss a deadline and will be fined a large amount if my accountant didn't get the message, the belief evaporates, and I ask for an acknowledgment. An important ingredient of the coordinated attack scenario, discussed below in Section 6.1, is enhanced risk.

Also, there are cases of *additive risk sensitivity*, in which a belief is created, not because of any relevant evidence, but because of adjusted assessments of risk. To continue the above example, suppose I now learn that the deadline has been extended. The belief that the message was received springs to life again, and I don't bother to ask for confirmation.[5] Such phenomena may seem more plausible to some readers if we recall that we are talking about practical belief—suppositions created in a deliberative context and on which we are willing to act in that context.

Additive adjustments to belief can depend on adjustments to risk tolerance, as well as on new evidence. The following example is from [30].

> Consider a nervous driver at a stop sign at a busy intersection on a dark night. He needs to drive across the intersection. He looks left. A car zooms by from that direction. He looks right. It's clear. He looks left, it's clear. But wait—he can't see what's going on to the right, and doesn't believe it's clear anymore. So he looks right. He repeats the process until he realizes that he'll never get across this way. Time is pressing. But he can't move unless the road is clear. So he lowers his standards, saying to himself "If it was clear to the right a second ago it's clear now." And he hits the gas.

---

[5]I realize that there is an alternative explanation of this example, in terms of probability and expected utility. But that is beside the point. People engage in belief-based practical reasoning far more commonly than utility calculations. If you like, you can think of risk-sensitive belief as a qualitative way of taking expected utility into account when doing reasoning of that sort.

This passage continues: "Sometimes, of course, there may be no intention to cross the intersection, and no belief—just a sort of desperate hope." People can act—perhaps out of desperation or frustration—without a supporting belief, simply in the hope that the action will have the desired outcome. And occasionally people may act like good Bayesian decision makers, acting on judgments about probability rather than on beliefs.

It may be difficult in practice to tell whether an agent was acting on a hope, a belief, or a probability, particularly since after the fact people tend to rationalize their decisions. But there are plausible examples of all three of these decision-making mechanisms.

4.   Achieving commonality

Attitudes held commonly by a group are iterated versions of individual attitudes. The formation of a common attitude can be illustrated with the simple case of a public announcement. In this case, the members of the group are presented with an announcement. Each is member is sufficiently familiar with the others to know they are capable of understanding the language of the announcement, and each member can observe not only the announcement but the others observing it.

Authors like David Lewis and Steven Schiffer formalize the inference of common attitudes in similar cases by showing how a finite set of axioms can guarantee commonality. This idea comes close to a solution, but falls short in at least two respects: (1) it doesn't explicitly address the defeasibility and even fragility of the assumptions that support commonality conclusions, or issues having to do with common belief revision and (2) it addresses cases where commonality follows from shared perceptions of the environment but has little to say about other cases.

Problem (1) can be addressed by using a nonmonotonic epistemic logic to formalize the construction of mutual attitudes. This is done in [27, 28], and explained with more attention to philosophical issues in [29].

The following axiom exemplifies the idea:

$$([a_1]\phi \land \neg Ab(a_1, a_i, \phi)) \rightarrow [a][a_i]\phi.$$

This axiom says that—assuming that the proposition p expressed by ϕ is not abnormal in the relevant way—if agent A_1 supposes[6] p then A_1 also supposes that agent A_i supposes p. If such an axiom is adopted for all formulas ϕ having the form $[a_1] \ldots [a_n]\psi$, where $a_j \in \mathcal{G}$, and all a_i, then $[a_i][\mathcal{G}]\psi$ is implied if there are no relevant abnormalities, where $[\mathcal{G}]$ is the common attitude induced by $[\,]$ and the group $\mathcal{G}$.

The result is not quite commonality, but something more attainable and just as good—that if there are no abnormalities for ϕ, $[a_j]\phi$, ..., then agent A_1 will *suppose* that the proposition p expressed by ϕ is a common supposition. This is a nonmonotonic logic, so abnormalities will be minimized—they will be assumed false unless there is some reason to suppose otherwise. Without such reasons, abnormalities do not need to be examined.

Ad hoc attitudes provide a simpler formalization. The members of a group may construct an *ad hoc* G-supposition attitude, expecting it to be common for the purposes at hand. An agent will not G-suppose a proposition unless there is reason to think that it is G-supposed in common. The following axiom is appropriate for this attitude.

$$\forall x \forall y \in \mathcal{G} \, \forall p (([x]p \wedge \neg Ab(x,y,p)) \rightarrow [x][y]p).$$

In particular applications, such axioms would need to be supplemented by an abnormality theory of the sort described in [12]; such a theory would also provide guidance about the revision of common belief in the presence of new evidence.

In the case of communicative suppositions, for instance, the relevant attitude would be conversational common ground (or supposition for the sake of the conversation), and cues indicating that an interlocutor hasn't heard or understood an utterance would count as abnormalities.

Problem (2) can be addressed by combining ideas of Herbert Clark and his co-authors with G-attitudes. [4] distinguishes personal and communal sources for common ground. Personal sources include features of the common environment. Communal sources use information about shared social background. For instance, speakers will assume knowledge of the nearby geography when talking to others from the same locality, and professional information when talking to someone in the same line of work,

---

[6]I am using 'suppose' here as a placeholder for whatever attitude is appropriate.

This can be formalized by assuming that declarative memory is not just a collection of stored propositions, but includes metadata, and in particular information about the circumstances in which an item was learned. A modality is a set of propositions. In possible worlds semantics, this would be a set of sets of worlds, but for our present purposes we only need to notice that metadata features classify propositions, so that a boolean combination of these features will determine a set of propositions, i.e. a modality.

For instance, if I was raised in Chicago, my memory may include a feature *Chicago-Native*, indicating that I learned it as a native of Chicago. If I'm an opera fan, it might also include a feature Opera-Fan. Then if I begin a conversation with a stranger, knowing she is a native of Chicago and an opera fan, I can use these two features to define a new modality, ⟦Chicago-Native & Opera-Fan⟧, and use this to initialize a conversation.

The idea that social institutions, as well as shared environment, can be a source of common attitudes can of course be applied in other domains. For instance, it can be used to explain the commonality of the rules of a game, of the regulations governing a market, and of social conventions.

5. Belief and decision-making

Although Lederman explicitly considers both common knowledge and belief, I myself, as I explained in Section 3, will be concerned only with belief, on the understanding that this is belief for the sake of some particular decision. We avoid having to deal with largely irrelevant philosophical issues about knowledge skepticism by confining our attention to belief.

Issues having to do with probability are yet another distraction. Sometimes people gamble, basing an action on probabilities. When an agent plays the odds, her actions rest on the hope, rather than the belief, that the outcome will be favorable. To the extent that the probability judgments are sound, such hopes will be well founded. But even the race-track gambler's decision is in part belief-based—she takes it for granted, among other things, that the track will pay for a winning ticket.

Belief is a matter of excluding possibilities, of taking some things for granted in a decision-making context. Which possibilities may sensibly and safely be excluded depends on the context, and in particular on the purposes to which the beliefs are to be put. We may exclude possibilities for various reasons: because we deem them irrelevant, or because it would be

complicated and time-consuming to take them into account, or because we think it does no harm to exclude them, or even because they are unpleasant or because we are too impatient to bother with them. Such reasons have little to do with the ideal rationality of classical economics. Some may be deliberative fallacies, but others are hard to classify. A factor like frustration doesn't seem rational, but there are times when it can be useful to act out of frustration, if that is the only way to escape a deliberative quandary.

Very frequently in daily life, when appropriate beliefs are deployed in a practical context, the need for probability is eliminated because nothing is left to chance. For instance, when someone who regularly drives to work plans her day, she automatically believes her car will start when she turn on the ignition. She doesn't bother to calculate the probability of failure. And she makes many other similar assumptions.

According to this picture of practical beliefs, they will have unit probability. This is the approach taken, for instance, in [31]. Of course, this idea makes sense only if both beliefs and probability spaces are relativized to the decision-making context. And of course it relies on the availability of efficient methods for choosing the alternatives that are appropriate for a given decision problem.

Since common beliefs are beliefs, they too will have unit probability. Such common beliefs must have been constructed constructed independently of the deployment of probabilities; e.g., in the process of framing a decision problem.

Some authors, such as Stephen Morris and Cédric Paternotte, propose a probabilistic account of common belief, based on the idea that belief is a matter of high, but not necessarily unit probability; see [14, 17]. This conception of common belief, and of belief in general, belongs to an entirely different approach to decision making, differing fundamentally from the one I've just sketched. This is one of several points at which more general issues in epistemology affect the treatment of common attitudes. I myself doubt that a successful probabilistic account of common attitudes can even be developed. If it can, I don't know whether it could be defended against skeptical arguments.

6.   Lederman's worries

Lederman raises two main objections to the use of common attitudes in
theories of cooperative activity: [10] argues that the assumption of common
knowledge (or belief) yields paradoxical results about two scenarios that
have been discussed in the technical literature. And [9] argues for the more
radical conclusion that it is impossible to achieve common knowledge and
belief about perceptual matters, and indeed about any substantive claim.
Both papers concentrate on *practical attitudes*—on knowledge and belief
as they figure in decision making.

6.1.   Coordinated attack

The coordinated attack scenario involves generals A_1 and A_2 who can only
communicate by sending insecure messages across enemy lines, who will
win (or are very likely to win) if they both attack, and will lose (or are
very likely to lose) if only one attacks. Its main purpose is to illustrate the
impossibility of obtaining common knowledge by message passing; but it's
plausible that similar situations can occur in real life.

An analysis of this situation in terms of expected utility is rendered
problematic by the difficulty A_1 will have in arriving at a probability that
A_2 will attack. Suppose for simplicity that the agents are identical decision-
making twins, and know this. Then A_1 can determine the probability that
A_2 will attack by imitating A_2's reasoning, and asking if in A_2's circum-
stances, A_1 would attack. This depends on A_2's probability that A_1 will
attack—but this is A_1's probability that A_2 will attack, which is what A_1
is trying to estimate in the first place. Fortunately, we do not have to ad-
dress this problem, because Lederman takes a belief-based approach to the
scenario, without making explicit assumptions about the relation between
belief and probability.

In discussing the example, Lederman invokes a characterization of com-
monsense rationality according to which a rational action "makes sense" or
is "explainable." But explainability is not at all the same thing as rational-
ity: irrational factors such as wishful thinking and carelessness can be used
to explain naturally occurring human decisions, but are hardly rational. I
am not confident that there is a robust and useful commonsense notion of
rationality, but if we wish to appeal to this notion we will need a better
characterization. For present purposes, I'll assume that an action in a set

of alternatives is CS-rational if in commonsense terms it is about as good as any alternative action.

Lederman takes it to be CS-rational for A_1 to attack after a finite number of messages have passed, but he doesn't say what the least such finite number is. And this is a problem, because it is hard to see how he or anyone could fix such a number. In fact, risk is a crucial feature of the CA scenario. We can't begin to say what it would be sensible for the generals to do without an idea of of the risks at stake in their decisions—but Lederman ignores risk entirely in his formulation of the scenario.

For the sake of definiteness, assume in what follows that it is about as likely for a message to fail as for it to succeed.

If losing the battle would be catastrophic, while winning would merely be moderately good, it seems pretty clear that doing nothing at all is CS-optimal for both generals. But if losing the battle would be a minor inconvenience, whereas winning would be outstandingly good, then attacking without sending any messages at all seems to be CS-optimal.

Neither of these extreme cases raises special problems about common belief. Consider, then, an intermediate case where risk is significant but not overwhelming, and assume it was CS-rational for A_1 to send a first message rather than to attack without sending a message. Then it doesn't seem as if any number of subsequent messages can produce a situation where it would be CS-rational for A_1 to attack. By assumption, it is not CS-rational for A_1 to attack after sending 0 messages. But after sending n messages, A_1's decision-making context looks the same as it does after 0 messages, because there is no more reason now than at first to believe that A_2 will attack.

What does this show? Is it a paradox? No: it merely shows that, in some versions of the CA scenario—for some values of likelihood and utility—the agents will be perplexed about what to do. Since people do occasionally find themselves in real quandaries—especially in war—this is neither surprising nor paradoxical.

Lederman's intuition that it is CS-rational for A_1 to attack after, say, 13 messages, and the experimental results to which he alludes, in fact have nothing in particular to do with CS-rationality or common attitudes, but with what we can expect of human agents who find themselves in quandaries. Consider the dithering driver scenario that was mentioned in Section 3. After several iterations, it's plausible to think that many human agents will pull out into the intersection, either hoping not to crash or (for

no rational reason) choosing to believe that no car is approaching on the blind side. Others might decide to turn around and try another route.

Some factors that are appropriate in the commonsense evaluation of decisions have little to do with rationality. Consider, for instance, the trade-off between sticking to previously made plans and willingness to abandon those plans to take advantage of apparent opportunities. We can recognize a spectrum of approaches, corresponding to obstinate and opportunistic personalities. We are familiar with individual differences along this spectrum, but it doesn't seem helpful to critique these differences in terms of rationality. Much the same can be said for impatience.

The dithering driver's decision has nothing to do with any useful conception of rationality, and everything to do with impatience and frustration in a difficult decision context. Though of course emotions and frustration can play a role in decision-making behavior, we're disposed to set aside such influences when we speak of rationality, even CS-rationality. The same considerations apply to the coordinated attack scenario, and it is hasty at best to conclude from human behavior in these cases that CS-rationality does not require mutual belief.

Even if there were a sensible policy for the CA scenario that would recommend attacking after, say, 2 successful iterations of message and acknowledgment, this would not show that mutual belief isn't required for coordinated action. This is because, as we mentioned in Section 3, belief thresholds can be adjusted in the course of deliberation, allowing the beliefs that support a choice to spring into existence.

Perhaps the distracted driver, after two left-right iterations, pulls out in the hopeful belief (belief for the sake of the decision) that a car isn't speeding toward him from the blind side. Perhaps the general who attacks after one or more messages does so in the hopeful belief that mutual belief has been achieved.

6.2.   Rubinstein's electronic mail game

[10] also discusses the electronic mail game, a scenario due to [18]. Rubinstein provides a game-theoretic formulation of the scenario, so that—unlike informal presentations of the coordinated attack problem—a quantitative formal analysis is available. But in other important respects, the electronic mail game is like the coordinated attack scenario.

Rubinstein showed that the strategy recommended by game theory in his example is independent of the number of messages that have been passed: no amount of message passing can affect the decisions of a game-theoretic agent. Lederman's point about this much-discussed example is only that the game-theoretic results don't match intuitions about CS-rationality.

It is hard to see what to make of this, because the relevant intuitions are far from robust. I myself do not think that Rubinstein's example raises any issues that differ significantly from versions of the coordinated attack scenario with precisely specified utilities and probabilities. Like the message-passing generals, humans who play this game and begin to pass messages will become frustrated after a while, and make a choice. But it isn't clear that these choices are conditioned by anything that could be called rational. I would claim that again, Lederman is misapplying rational criteria to what agents in a hard and perplexing decision context and influenced by human emotions might be expected to do.

If I suppress emotional factors, by imagining that the agents in this scenario are utility-optimizing computer programs, Rubinstein's result doesn't strike me as counterintuitive.

7.   Lederman's sailboat scenario

[9] uses attitudes about the value ranges of continuous quantities to argue for the much stronger conclusion that our minds are not in fact "open to each other," which "casts doubt on the idea that people ever have common knowledge or its relatives." As I said, I am not concerned here with knowledge. But Lederman thinks that belief is susceptible to similar arguments.

Lederman's scenario has nothing important to do with sailboats. Two players, visible to each other, observe a long, thin object—the "mast." The mast is then replaced with another that has a randomly selected length (presumably, within a range fairly close to the first mast). Each agent is assigned a button, which they must then decide whether to press. If the new mast is taller than the old one, the reward is +1000. If it is not taller, and neither player presses a button, the reward is 0. Otherwise the reward is -100. On the classical account, and using belief as the relevant attitude, a player A_1 should press the button if and only if A_1 believes that it's mutually believed that the new mast is taller.

Lederman presents his argument using knowledge. A version of it for belief runs as follows.

1. The masts will look to have a certain height to the players; Lederman uses "looks r centimeters high to A," formally LOOKS(A,r), in a slightly peculiar way according to which, for instance, LOOKS(A,100) and LOOKS(A,100.01) are consistent—that is, LOOKS(A,x) will be true over an interval (r_1, r_2). This interval will have a midpoint, say r. Let LOOKS′(A, r, ρ) hold iff r is the midpoint of the open interval with radius ρ over which LOOKS(A,x) holds; then there will be at most one x such that LOOKS′(A, x, ρ) is true.

2. Interpersonal estimates of perceptions about continuous quantities have a margin of error. Suppose with Lederman that for estimating the height of the mast this is .03, that the margin is known by the agents, and—crucially—that it is fixed at .03 in all epistemically accessible worlds. In terms of agent beliefs, this means that LOOKS′(A_1, x, ρ) holds, then so does ‹A_1›LOOKS′(A_2, y, ρ), for all y such that $|x - y| \leq .03x$. (Here, angle brackets are labeled possibility signs or diamonds, so that ‹A_1› stands for "for all A_1 believes.") And likewise for A_2's opinions about how things look to A_1. Not only do these things hold, but both agents are aware that they hold.

3. In terms of the possible worlds semantics for belief, this means that if LOOKS′$(A_1, 100, \rho)$ holds in world w then for all x such that .97 ≤ x ≤ 1.03, there is a world w(A_1,x) that is A_1-accessible from w, such that LOOKS′(A_2, x, ρ) holds in w(A_1,x); the range of such worlds is the interval $I_1 = (.97, 1.03)$ with diameter .06. For similar reasons, the interval of worlds A_2-accessible from worlds in I_1 in which LOOKS′$(A_1, 100, \rho)$ holds is approximately $I_2 = (.941, 1.061)$, with diameter .12. In general, each iteration of the sequence
   ‹A_1›‹A_2›‹A_1›‹A_2› ... LOOKS(A_2, k)
increases the diameter of the error interval by more than .06. Lederman concludes that the agents can't commonly believe that the height is greater than any positive height ϵ, because there will be a point in the attitude iteration including a world where the error interval includes ϵ.

If we accept this sort of argument, we may get more than we bargained for, A similar argument would show that if the mast looks to have height r for a single agent A_1, then it must look to A_1 to have any other height. If the mast looks to A_1 to have height x, then for some fixed ϵ, less than the perceptual threshold for height, it looks to A_1 to have heights $x \pm \epsilon$. By iterating this argument, it looks to have any height whatever. This in fact is Michael Dummett's perceptual version of the sorites paradox, [5].

This similarity to a paradox makes Lederman's argument dubious. So I will consider an improved version, suggested by a referee of this paper.

> Two agents A_2 and A_3 are told that each will be assigned an integer greater than 1, and that the integers will be consecutive. A_2 is given the number 2 and A_3 the number 3. Consider a set $\{w_{i,j} \mid i,j > 0\}$ of alternatives, with the understanding that in $w_{i,j}$ A_2 is given i and A_3 is given j, and let R_2 and R_3 be the epistemic accessibility relations for A_2's and A_3's beliefs, respectively. Then $w_0 R_2 w_1 R_3 w_2 \ldots R_3 w_n$ iff $w_{2n} = w_{2,i}$ for some $i \in [1, n+3]$, and likewise $w_0 R_3 w_1 R_2 w_2 \ldots R_2 w_n$ iff $w_{2n} = w_{2,i}$ for some $i \in [1, n+3]$.
>
> No matter how large n may be, it seems that the agents can't have a common belief that n bounds both agents' numbers.
>
> We can make this practical by probing the agents with a positive integer n and asking them to press a button, with rewards and penalties like those of the sailboat example.

This scenario doesn't use continuous quantities and is not at all similar to the sorites paradox. But now is very like the coordinated attack problem and can be treated in much the same way. It presents the contestants with a practical problem that may well have no solution in terms of economic rationality. But people have other ways of making decisions, some better than others. In this case, a participant may arrive at the belief that, say, it's commonly believed that 10 is an upper bound by eliminating some alternatives. This could be because 10 is a salient number, or it could be a matter of jumping to a conclusion for no very good reason. Or a participant might press a button as a hopeful gamble, without forming any relevant belief at all.

8. Conclusion

Early skeptical doubts about common attitudes emerged in the 1980s. These can be addressed by concentrating on common belief and using nonmonotonic logics to respond to the technical challenge of accounting for how agents can arrive at and reason with these attitudes. Lederman's more recent doubts about the appropriateness of common attitudes hinge on idiosyncratic intuitions about "commonsense rationality," which don't provide clear guidance when applied to the behavior of human agents in challenging circumstances.

It is less interesting then, to confront Lederman's conclusions with opposing intuitions about rationality than to ask if common attitudes such as conversational common ground and the common beliefs at stake in a market or a game can be situated within a sensible theoretical approach to practical reasoning. I have argued that a framework based on defaults and rough estimates of likelihoods and risks can account for how common beliefs can originate and how, like other beliefs, they can play a part in decision making. On this picture, common beliefs are readily inferrable, not by any extraordinary and unusual feat of reasoning, but by methods that are constantly in play in our everyday life.

Probably the main source for skepticism about common attitudes is a misunderstanding about the scope and proper place of economic rationality in deliberative contexts. Calculation of rational optima makes good sense when a problem can be framed in terms that enable such calculations to be made. But in real life we are often confronted with problems that can't be framed that way—and this includes the problem of modeling a decision problem. In these cases commonsense reasoning mechanisms such as rough assessments, intuitions about relevance, and defaulr reasoning can come into play. And these mechanisms can support common attitudes—for instance, by justifying the assumption that public announcements create common beliefs.

But this response to philosophical doubts about common attitudes—however successful—doesn't suffice to justify invoking these attitudes in game theory, protocol analysis, the theory of conversations, any of the other areas where they may seem theoretically appropriate. That has to be done on a case-by-case basis, using the methods of the relevant discipline. For theoreticians who postulate common attitudes, then, the philosophical

part of this paper is not directly relevant, although the more technical ideas in [28] about the reasoning that supports common attitudes may be useful.

References

[1] R. J. Aumann, *Agreeing to Disagree*, **Annals of Statistics**, vol. 4(6) (1976), pp. 1236–1239, DOI: https://doi.org/10.1214/aos/1176343654.

[2] S. Bucheli, R. Kuznets, T. Studer, *Justifications for Common Knowledge*, **Journal of Applied Non-Classical Logics**, vol. 21(1) (2011), pp. 35–60, DOI: https://doi.org/10.3166/jancl.21.35-60.

[3] H. H. Clark, C. R. Marshall, *Definite Reference and Mutual Knowledge*, [in:] A. Joshi, B. Webber, I. Sag (eds.), **Elements of Discourse Understanding**, Cambridge University Press, Cambridge, England (1981), pp. 10–63.

[4] H. H. Clark, M. Schober, *Understanding by Addressees and Overhearers*, **Cognitive Psychology**, vol. 21(2) (1989), pp. 211–232.

[5] M. A. Dummett, *Wang's Paradox*, [in:] R. Keefe, P. Smith (eds.), **Vagueness: A Reader**, The MIT Press, Cambridge, Massachusetts (1997), pp. 99–118.

[6] J. Geanakoplos, *Common Knowledge*, [in:] R. Aumann, S. Hart (eds.), **Handbook of Game Theory, with Economic Applications**, chap. 40, vol. 2, North-Holland, Amsterdam (1994), pp. 1437–1496, DOI: https://doi.org/10.1016/S1574-0005(05)80072-4.

[7] P. N. Johnson-Laird, *Mutual Ignorance: Comments on Clark and Carlson's Paper*, [in:] **Mutual Knowledge**, Academic Press, London (1982), pp. 40–45.

[8] H. Lederman, *Common Knowledge*, [in:] M. Jankovic, K. Ludwig (eds.), **The Routledge Handbook of Collective Intentionality**, Routledge, New York (2017), pp. 181–195.

[9] H. Lederman, *Uncommon Knowledge*, **Mind**, vol. 127(508) (2017), pp. 1069–1105, DOI: https://doi.org/10.1093/mind/fzw072.

[10] H. Lederman, *Two Paradoxes of Common Knowledge: Coordinated Attack and Electronic Mail*, **Noûs**, vol. 52(4) (2018), pp. 921–945, DOI: https://doi.org/10.1111/nous.12186.

[11] D. K. Lewis, **Convention: A Philosophical Study**, Harvard University Press, Cambridge, Massachusetts (1969), DOI: https://doi.org/10.1002/9780470693711.

[12] V. Lifschitz, *Circumscriptive Theories: A Logic-Based Framework for Knowledge Representation*, **Journal of Philosophical Logic**, vol. 17(3) (1988), pp. 391–441, DOI: https://doi.org/10.1007/BF00297512.

[13] M. Merritt, G. Taubenfeld, *Knowledge in Shared Memory Systems*, **Distributed Computing Archive**, vol. 7 (1991), pp. 99–109, DOI: https://doi.org/10.1007/BF02280839.

[14] S. Morris, H. S. Shin, *Approximate Common Knowledge and Co-ordination: Recent Lessons from Game Theory*, **Journal of Logic, Language, and Information**, vol. 6(2) (1997), pp. 171–190, DOI: https://doi.org/10.1023/A:1008270519000.

[15] D. G. Novick, K. Ward, *Mutual Beliefs of Multiple Conversants: a Computational Model of Collaboration in Air Traffic Control*, [in:] R. Fikes, W. Lehnert (eds.), **Proceedings of the Eleventh National Conference on Artificial Intelligence**, American Association for Artificial Intelligence, AAAI Press, Menlo Park, California (1993), pp. 196–201.

[16] P. Panangaden, K. Taylor, *Concurrent Common Knowledge: Defining Agreement for Asynchronous Systems*, **Distributed Computing**, vol. 6(2) (1992), pp. 73–94, DOI: https://doi.org/10.1007/BF02252679.

[17] C. Paternotte, *The Fragility of Common Knowledge*, **Erkenntnis**, vol. 82 (2017), pp. 451–472, DOI: https://doi.org/10.1007/s10670-016-9828-4.

[18] A. Rubinstein, *The Electronic Mail Game: Strategic Behavior Under 'Almost Common Knowledge'*, **American Economic Review**, vol. 79(3) (1989), pp. 385–391.

[19] S. Schiffer, **Meaning**, Oxford University Press, Oxford (1972).

[20] S. Schiffer, **Remnants of Meaning**, The MIT Press, Cambridge, Massachusetts (1987).

[21] H. A. Simon, *Rational Choice and the Structure of the Environment*, **Psychologicial Review**, vol. 63(2) (1956), pp. 129–138, DOI: https://doi.org/10.1037/h0042769.

[22] D. Sperber, D. Wilson, *Mutual Knowledge and Relevance in Theories of Comprehension*, [in:] N. Smith (ed.), **Mutual Knowledge**, Academic Press, London (1982), pp. 61–85.

[23] D. Sperber, D. Wilson, **Relevance: Communication and Cognition**, 1st ed., Harvard University Press, Cambridge, Massachusetts (1986).

[24] D. Sperber, D. Wilson, **Relevance: Communication and Cognition**, 2nd ed., Blackwell, Oxford (1995).

[25] R. Stalnaker, *Common Ground*, **Linguistics and Philosophy**, vol. 25(5–6) (2002), pp. 701–721, DOI: https://doi.org/10.1023%2FA%3A1020867916902.

[26] R. H. Thomason, *The Multiplicity of Belief and Desire*, [in:] M. P. Georgeff, A. Lansky (eds.), **Reasoning about Actions and Plans**, Morgan Kaufmann, Los Altos, California (1987), pp. 341–360.

[27] R. H. Thomason, *Propagating Epistemic Coordination Through Mutual Defaults I*, [in:] R. Parikh (ed.), **Theoretical Aspects of Reasoning about Knowledge: Proceedings of the Third Conference (TARK 1990)**, Morgan Kaufmann, Los Altos, California (1990), pp. 29–39, DOI: https://doi.org/10.5555/1027014.1027022.

[28] R. H. Thomason, *Modeling the Beliefs of Other Agents*, [in:] J. Minker (ed.), **Logic-Based Artificial Intelligence**, Kluwer Academic Publishers, Dordrecht (2000), pp. 375–473, DOI: https://doi.org/10.1007/978-1-4615-1567-8_17.

[29] R. H. Thomason, *The Beliefs of Other Agents* (2001), http://web.eecs.umich.edu/~rthomaso/documents/nmk/otheragents.pdf.

[30] R. H. Thomason, *Belief, Intention, and Practicality: Loosening Up Agents and Their Propositional Attitudes*, [in:] F. Lihoreau, M. Rebuschi (eds.), **Epistemology, Context, Formalism**, Springer Verlag, Berlin (2013), pp. 167–184, DOI: https://doi.org/10.1007/978-3-319-02943-6_10.

[31] B. C. van Fraassen, *Representation of Conditional Probabilities*, **Journal of Philosophical Logic**, vol. 5(3) (1976), pp. 417–430, DOI: https://doi.org/10.1007/BF00649400.

Richmond H. Thomason

University of Michigan
Philosophy Department
2251 Angell Hall
48109-1003 Ann Arbor
Michigan, USA

e-mail: rthomaso@umich.edu