The Role of Questions, Circumstances, and Algorithms in Belief

Jens Kipper[1], Alexander W. Kocurek[2], and Zeynep Soysal[1]

[1] University of Rochester
[2] Cornell University

A recent approach to the problem of logical omniscience holds that belief is *question-sensitive*: belief is not a two-place relation between an agent and a proposition, but rather a three-place relation between an agent, a question, and an answer to that question (Pérez Carballo, 2016; Yalcin, 2018; Hoek, 2022). While the question-sensitive approach can avoid some logical omniscience problems, we argue that it suffers from nearby problems. Recognizing the flaws of this approach, however, naturally leads to a more promising solution.

On the classical picture, a belief state is modeled as set of possible worlds, viz., those worlds compatible with what the agent believes (Hintikka, 1962; Stalnaker, 1984).[1] An agent believes a proposition P iff their belief state B entails P (i.e., $B \subseteq P$). Commonly, this picture is combined with the idea that there are constitutive links between belief and (rational) action. An agent who believes P is disposed to act on P, which means, roughly, that they are disposed to bring about desire-satisfaction in worlds in which P and the agent's other beliefs are true (see Stalnaker 1984). This picture of belief has many advantages (Stalnaker, 1984; Yalcin, 2018). For one, it explains the holistic nature of belief update. If an agent comes to believe they left their keys on the table, they will often automatically come to believe their keys are on the table, their keys are not in their pocket, and so on. It also explains how we can attribute to agents beliefs in propositions they haven't considered. If a driver sees a moose in the road and swerves to avoid a collision, it is fair to say they believe the moose is larger than a golf ball, even if they have not explicitly articulated this belief.

Despite these advantages, however, the classical picture faces the well-known *problem of logical omniscience:* it incorrectly predicts that belief is closed under entailment. For if a belief state B entails P and P entails Q, then B entails Q. So for example, on the classical picture, since 107 is necessarily prime, if Chip believes the CVV of his debit card is 107, he must also believe the CVV is prime. Yet it is clear that Chip can have the former belief without the latter. As a special case, the classical picture predicts that an agent always believes any necessary truth and that their beliefs are closed under equivalence. So, continuing the previous example, the classical account predicts that Chip believes 107 is prime and that if he believes his CVV is 107, then he believes his CVV is the 28th prime number.

The question-sensitive account of belief is designed to retain the core benefits of the classical account while allowing for failures of closure. On this picture, agents do not simply believe propositions: they believe *answers to questions*. Questions effectively determine which propositions are "visible" to an agent: relative to a question, agents can only believe propositions that are answers to that question. So, while Chip may believe his CVV is 107 relative to the question *What's the CVV?*, he may not believe this relative to the question *What information does my debit card contain?* (e.g., he may not be able to recall his debit card includes the CVV, even though he could recite the CVV if asked about it explicitly).

[1]To account for egocentric belief, i.e., belief about one's location in logical space, belief states can also be modeled as sets of *centered* worlds, where a centered world is a pair of a possible world and an individual at a time—see Lewis 1979. While we find this refinement of the classical picture useful, we still construe the worlds involved as uncentered possible worlds here, for the sake of simplicity.

Most defenders of this account follow the early literature on the semantics of questions, which model questions as partitions on possible worlds (Hamblin, 1973; Groenendijk and Stokhof, 1984; Lewis, 1988). A belief state is thus a set of *pointed partitions*, i.e., pairs $\langle Q, A\rangle$ where Q is a partition on possible worlds (a question) and A is a union of Q-cells (an answer to Q).[2] An agent believes A relative to Q iff (i) A is an answer to Q, and (ii) there is a pointed partition $\langle Q, B\rangle$ in their belief state such that $B \subseteq A$. This account retains the main benefits of the classical account (e.g., the holistic nature of belief update) while allowing for failures of closure. For example, Chip may believe the CVV is 107 relative to *What's the CVV?* without believing the CVV is prime relative to *Is the CVV prime?*, since these are different questions.[3]

There are two main problems with such accounts. First, they all validate the following implausible closure principle: if A entails B, both of which are answers to Q, then if an agent believes A relative to Q, they also believe B relative to Q.[4] For example, the proposition *The CVV is prime* is an answer (albeit a partial one) to *What's the CVV?*, and it is entailed by *The CVV is 107*. Hence, these accounts implausibly predict that if Chip believes the CVV is 107 relative to *What's the CVV?*, then he also believes the CVV is prime relative to the same question.[5] This closure principle also entails two weaker principles, both of which are problematic. The first involves necessary answers: if an agent has any belief relative to Q, then they believe every necessarily true answer to Q. Thus, if Chip believes 107 is either prime or composite relative to *Is 107 prime?*, then he also believes 107 is prime relative to this question. The second problematic principle involves equivalence: if an agent believes A relative to Q, where A is necessarily equivalent to B and Q is necessarily equivalent to R (meaning a complete answer to one is a complete answer to the other), then they believe B relative to R. Thus, if Chip believes his CVV is among 101, 103, 107, ... (listing all the 3-digit primes) relative to *Is the CVV among 101, 103, 107, ...?*, then he also believes his CVV is a 3-digit prime relative to *Is the CVV a 3-digit prime?*. These predictions seem as problematic as the predictions of the classical account.

Second, question-sensitivity by itself is not suitable for explaining many kinds of failures of logical omniscience. Question-sensitive accounts are supposed to explain failures of logical omniscience in terms of limitations of computational resources. For instance, Yalcin (2018) argues that failures to recognize the logical consequences of our beliefs are due to our inability to resolve logical space in full detail.[6] A partition represents how finely an agent resolves logical space. Finer resolutions reduce logical ignorance but come with higher "encoding costs" in the form of *information potential*—roughly, the number of propositions needed to construct the partition, which is just the logarithm (base 2) of the number of cells in the partition (bits) rounded up to the nearest integer. Thus, failures of logical omniscience are connected to an agent's lack of resources for representing very fine-grained partitions.

[2] Some authors impose further constraints on belief states. See, e.g., Yalcin 2021; Hoek 2022. Our criticisms of question-sensitive accounts is independent of such constraints.

[3] For a question-sensitive account of the link between belief and action, see Hoek (2022).

[4] Yalcin (2018) refers to this principle (for "accessible belief") as closure under "visible consequence". While he acknowledges this principle may be undesirable, his reasons are different: his worry is that it renders too many propositions visible to the agent. Even if we set aside these sorts of worries, however, we think this closure principle has undesirable consequences.

[5] Teague (forthcoming) gives two counterexamples to a stronger principle: if every complete answer to Q entails a complete answer to R, and an answer A to Q entails an answer B to R, then if an agent believes A relative to Q, they believe B relative to R. However, while some accounts validate this principle (Hoek, 2022, forthcoming), others do not (Yalcin, 2018, 2021). Moreover, Teague's counterexamples rely on lacking concepts and lacking attention. By contrast, our counterexample does not: Chip may possess the concept PRIME and be attending to the question of whether the CVV is prime yet still fail to believe it is. Our counterexample is thus more pressing for question-sensitive accounts.

[6] See Pérez Carballo (2016) and Hoek (forthcoming) for similar ideas.

However, there are other, and plausibly more relevant, sources of computational costs associated with determining logical consequences. Consider, e.g., the game of chess. The rules of chess determine, for any given position, whether there is a forced checkmate. Moreover, a competent player can resolve any chess position in full detail. They may even have an algorithm, such as minimax, that can decide whether there is a forced checkmate in any given position. Nevertheless, in many positions, this algorithm would take much too long to conclude, and thus chess players often fail to know (or have a belief about) whether there is a forced checkmate. By contrast, it is much easier to determine the position of the chess board. Even though the information potential of *Is there a forced checkmate?* is small (1 bit) compared to *What's the position?* ($\approx$133 bits), the relevant computational costs of the former relate to time-complexity, not "encoding costs". Moreover, there are other kinds of failures of logical omniscience that aren't captured by the question-sensitive approach. One comes from how a question is presented to an agent. For example, Chip may be able to recall his CVV as an answer to a multiple-choice question (with a "None of the above" option) but unable to recall it as an answer to a fill-in-the-blank question. Others come from an agent's psychological state, or from the external circumstances they are in. For instance, Chip may be able to recall his CVV in normal conditions but not when he is stressed or when loud music is playing in the background.

The account we suggest is designed to encompass these other sources of failures of logical omniscience as well. Like both the classical account and question-sensitive accounts, our account is based on the idea that an agent who believes ϕ is disposed to act on the information associated with ϕ. We use two implications of this idea.

First, as was explained above, a disposition to act on (the information associated with) ϕ can be understood, roughly, as one to bring about desire-satisfaction in worlds where it is true. Here, desire-satisfaction is the manifestation condition of this disposition and (the information associated with) ϕ is its trigger condition. Now, dispositions are generally dependent on *ceteris paribus* or normal conditions. These are the conditions that need to be in place for a disposition to be manifested, if the trigger occurs. For instance, a normal condition for a match to manifest its disposition to light if struck is the presence of oxygen. The same applies to doxastic dispositions. Accordingly, belief is not just *question*-sensitive but, more generally, *circumstance*-dependent, i.e., dependent on the set of normal conditions being considered. This implies that for an agent to be able to satisfy their desires, they not only need to have true beliefs, but the relevant circumstances must obtain as well. These circumstances can concern not just the question an agent faces, but anything about their cognitive or physiological functioning, or their external conditions.[7]

[7] In a similar vein, Elga and Rayo (2021) characterize the mental states of agents using "conditons". They develop a model of fragmented agents in terms of "access tables", which map conditions to credence functions, and then postulate that it is appropriate to attribute a full belief to a fragmented agent iff the dispositions predicted by that belief are sufficiently similar to those predicted by the access table. However, our approach is different in several respects. First, Elga and Rayo use conditions to characterize fragmentation: different conditions correspond to different fragments of an agent's mental state. By contrast, our use of circumstances is not for characterizing fragments but rather for characterizing the dispositions associated with a belief. Our model of belief does not require agents to be fragmented to avoid logical omniscience. Second, while Elga and Rayo reduce full belief to dispositions to act, they do not attempt to formally represent the dispositions associated with a given belief, whereas our model does. Third, Elga and Rayo postulate an "obviousness function" associated with each condition that determines (roughly) which inferences are "obvious" in that fragment and which are not. Our account instead appeals directly to an agent's dispositions (and their decision-making algorithms) to determine which inferences are they find "obvious" in a given circumstance. Lastly, on their account, each condition is associated with a normal credence function, which means that the information accessible relative to given condition is closed under entailment. Thus, their account validates something akin to the question-sensitive closure principle we reject.

Second, it is highly plausible that an agent's dispositions to act are produced by (internal) algorithms. Algorithms can thus play an important explanatory role in an account of belief: whether one acts on a piece of information in certain circumstances depends on the algorithm one is disposed to deploy in these circumstances.[8] On our view, then, believing that ϕ means being disposed to employ an algorithm which outputs acting on the information associated with ϕ in normal circumstances associated with ϕ. Typically, this means that the algorithm outputs the relevant action in a relatively short amount of time: since most of the decisions we face are time-sensitive, an agent whose algorithms are very slow will usually not be able to act in ways that bring about desire-satisfaction.[9]

We formalize this account as follows. Given a background language $\mathcal{L}$ and a set of worlds W, we capture circumstance-dependency by assuming that each $\phi \in \mathcal{L}$ at each $w \in W$ is associated with a set $C_{\phi,w} \subseteq W$ of "normal" worlds for ϕ at w.[10] Let $[\phi]$ be the set of worlds where ϕ is true. To capture algorithm-dependency, we first model decision problems as pairs $\langle O, t \rangle$, where O is a set of options or choices the agent can make and t is the amount of time an agent has to make their choice. Let Δ_w be the set of decision problems an agent faces in w. Then at each world $w \in W$, an agent is associated with a decision algorithm α_w that takes decision problems in Δ_w as inputs and outputs an option after a certain amount of time if it halts. We write $\alpha_w(O, t) = a$ if α outputs $a \in O$ by t; otherwise, $\alpha_w(O, t)$ is the "null-action" (do nothing).

There are many ways to flesh out the notion of "acting on" some information. For concreteness, we assume agents "act" on information by avoiding options that would be weakly dominated were the information accurate (see Hoek 2022). This is a natural way of making precise the idea that acting on the information that ϕ means acting in ways that bring about desire-satisfaction in ϕ-worlds. More precisely, where $P \subseteq W$, an option $a \in O$ is weakly P-dominated by $b \in O$ if the utility an agent receives by performing a is no greater, and in some cases strictly less, than the utility they receive by choosing b in any P-world. An agent acts on P, then, if they avoid weakly P-dominated options.

Now, it seems plausible that an agent can be associated with a belief that ϕ without having a disposition to act on $[\phi]$ itself. To take an extreme case, if $[\phi] = \varnothing$ (i.e., ϕ is impossible), then no option weakly $[\phi]$-dominates any other. Thus, if we allow belief in impossible propositions, the dispositions to act on them cannot just be the disposition to act on $[\phi]$ itself. However, even for contingent propositions, having the disposition to act on *that* proposition is plausibly not the same as believing it. Suppose the bank teller asks Chip for his CVV to verify his identity. Arguably, if Chip believes his CVV is 107, he should be disposed to answer "107" in this sort of case. But if the belief that his CVV is 107 is only associated with the disposition to act on [The CVV is 107], then we will fail to predict this disposition. For there are [The CVV is 107]-worlds where answering "107" does not bring about desire-satisfaction for Chip—e.g., the bank teller might be trying to steal his debit card information to commit fraud, and in those worlds, answering with a random false number would be preferable. Indeed, there are relatively few decision problems that have weakly dominated options at *all* [The CVV is 107]-worlds. But, intuitively, many of those worlds do not seem relevant for characterizing the information Chip is disposed to act on.

One way to resolve this issue is to say that a belief ascription is not associated with a disposition to act on the content of the ascribed belief, but rather to act on some related proposition.

[8] Algorithmic accounts of belief and knowledge have been developed by Parikh (1987); Halpern et al. (1994). See Soysal 2022 for a recent algorithmic account that is tied to a possible-worlds account of content.

[9] See Stalnaker 1991, 437 for discussion of this issue.

[10] Plausibly, it would be more adequate to construe circumstances in terms of centered worlds, accounting for the idea that these represent the circumstances an agent may find themselves in (see footnote 1.) Since this issue isn't relevant to our discussion, we stick with uncentered worlds here, for the sake of simplicity.

For the sake of generality, let us introduce a contextually-supplied function $(\cdot)$, which takes formulas as inputs and outputs propositions (cf. Williamson's (2009) use of "counterparts" of formulas for knowledge ascriptions). We can think of (ϕ) as the "actionable content" of ϕ for the agent, i.e., the relevant contextual information that an agent is disposed to act on given that it is correct to ascribe to them a belief that ϕ. For example, in the case of Chip, (The CVV is 107) may be the conjunction of the proposition that the CVV is 107 together with other contextual information. If $[\phi] = \varnothing$, then (ϕ) might be a metalinguistic proposition that the sentence ϕ is true (perhaps conjoined with relevant contextual information). Then our account can be stated as follows: an agent believes ϕ at w iff for all $v \in C_{\phi,w}$, all $\langle O, t\rangle \in \Delta_v$, and all $a \in O$, if a is weakly (ϕ)-dominated by some $b \in O$, then $\alpha_v(O, t) \neq a$.

This account is suitable for solving the problem of logical omniscience, since it does not entail any of the problematic closure conditions that the classical account or the question-sensitive accounts are committed to. Due to the fact that we construe belief as relative to circumstances, an agent's believing ϕ only entails that they believe ψ if it is both the case that $(\phi) \subseteq (\psi)$, i.e., the actionable content of ϕ contains that of ψ, and that $C_{\psi,w} \subseteq C_{\phi,w}$. The second conjunct ensures that if someone can bring about desire-satisfaction in all circumstances associated with ϕ, they can do so in all those associated with ψ. To see how this account avoids closure under entailment, assume that $(\phi) \subseteq (\psi)$ and that $C_{\psi,w} \nsubseteq C_{\phi,w}$. Then an agent who is disposed to employ an algorithm that outputs acting on the information that (ϕ) in $C_{\phi,w}$ but not in $C_{\psi,w}$ believes that ϕ without believing that ψ.

As an illustration, it might be part of the circumstances associated with a belief which task one is facing. For instance, to believe that the CVV is among 101, 103, 107, ..., one plausibly needs to have an algorithm that quickly outputs "Yes" if one is (verbally) asked "Is the CVV among 101, 103, 107, ...?" but not one that quickly outputs "Yes" if one is asked "Is the CVV a three-digit prime?". Since an agent can have the former kind of algorithm without having the latter, this explains why one can believe one without believing the other, despite the fact that the representational content of the two beliefs is the same, and despite the fact that the questions the agent are asked are equivalent.

This example is also suggestive of how one can account for failures to have beliefs whose contents are necessarily true. If (ϕ) is necessarily true, then the only reason why an agent can fail to believe ϕ is because they do not have an algorithm that outputs desire-satisfaction in $C_{\phi,w}$. Assume, e.g.—as seems plausible—that to believe that 107 is prime, one needs to be able to respond "Yes" in some limited amount of time if one is asked "Is 107 prime?", given that one desires to give the correct answer. On our account, this means that this task is among the circumstances associated with "107 is prime". Hence, since an agent can fail to have an algorithm that outputs the correct response sufficiently quickly, an agent can fail to believe that 107 is prime.

While we reject commonly held closure conditions on belief, our account can nevertheless capture the fact that beliefs often come in packages. For instance, an agent who believes that they just saw a moose larger than a golf ball will typically also believe that they just saw an animal, that they just saw an object larger than a bottle cap, etc. On our account, this can be explained by appealing to the algorithms involved. Often, one algorithm available to an agent grounds many of their dispositions to act and, thus, confers many beliefs to them. Furthermore, many beliefs are very commonly (but not necessarily) conferred by specific algorithms. Taken together, this explains why an agent with a particular belief can often be expected to have many other beliefs—due to the fact that the algorithm that confers one of these beliefs also confers the others. For instance, an algorithm that produces desire-satisfaction with respect to questions about whether some number is divisible by 6 usually also produces desire-satisfaction with respect to questions about whether some number is divisible by 2, because the standard method most people use to check divisibility by 6 involves checking for divisibility by 2. This can explain why someone who believes, say, that 294 is divisible by 6 will typically also believe that 294 is divisible by 2. But they won't necessarily believe the latter, for they might use alternative methods for answering questions about divisibility by 6, such as having memorized all numbers divisible by 6 up to 294. The holistic nature of belief update, mentioned above, can be explained in a similar manner: since possession of one algorithm often confers many different beliefs, we can expect an agent who acquires a new belief to automatically acquire many other beliefs. Finally, similar types of constraints on algorithms could be used to capture rational or competent belief. For instance, one might construe logically competent agent as those who employ algorithms that "respect" the standard inference rules for the logical connectives, in the sense that their algorithm computes, say, a conjunction to be true just in case it computes its conjuncts to be true.

The involvement of algorithms also allows us to explain many types of failure of logical omniscience. Very commonly, if an agent fails to believe a logical consequence of something they believe, this is either because they do not have an algorithm that allows them to derive this consequence, or because they only have an algorithm that would take too many computational resources to derive it. Take again Chip, who believes that his CVV is among 101, 103, 107, ...but does not believe that it is prime. The latter might be because he does not have an algorithm that can determine primality, or, more likely, because he does have such an algorithm, but it would take too long for him to run it. Either way, Chip is unable to give the correct answer to "Is the CVV prime?" in a reasonable amount of time and hence, he lacks the relevant belief.

In closing, let us return to the question suggested in the title: what role do questions, circumstances, and algorithms play in belief? We argued that question-sensitive accounts are limited in scope and only solve restricted versions of the problem of logical omniscience while leaving the full version unresolved. Even so, we argued that question-sensitive accounts are picking up on a specific instance of a more general feature of belief, viz., circumstance-dependency. By modeling belief as sensitive to both ceteris paribus conditions and the algorithms an agent uses to act, we developed a more robust and viable account of belief that solves the problem of logical omniscience in full generality.

Our account of belief leaves open some unanswered questions that we leave to future research. First, our discussion here has focused on full belief, and thus we have left it open how to generalize our account to partial belief or credences. Second, it might be useful to specify the notion of "actionable content" further by introducing constraints on the $(\cdot)$ function. Such constraints have ramifications for the logic of belief, the details of which could be examined in future work.[11] Lastly, as we explained above, our account reduces the rationality of belief to constraints on the decision-making algorithms an agent employs. We suggested one constraint we might impose on logically competent agents, i.e., agents who understand the logical connectives. We can obtain more idealized forms of rationality by imposing stronger constraints, e.g., requiring an agent's algorithm to avoid weakly dominated actions in *all* circumstances. We leave it open here how other aspects of rationality can be captured by introducing constraints on algorithms.

[11]One constraint worth considering is the following: if $[\phi] \subseteq [\psi]$, then $(\phi) \subseteq (\psi)$. This ensures that if ϕ entails ψ, then being disposed to act on ϕ in certain circumstances entails being disposed to act on ψ *in the same circumstances*. (Note: this does not imply that an agent who believes ϕ also believes ψ since the set of circumstances associated with ϕ and ψ might still differ.) Another plausible constraint: where T is a truth predicate, $(\phi) \subseteq [\phi] \cup [T(\ulcorner\phi\urcorner)]$. This would ensure that if $[\phi] = \varnothing$, then $(\phi) \subseteq [T(\ulcorner\phi\urcorner)]$, i.e., (ϕ) at least contains the metalinguistic information that the sentence ϕ is true.

References

Elga, Adam and Rayo, Agustín. 2021. "Fragmentation and Logical Omniscience." *Noûs* 1–26. doi:10.1111/nous.12381.

Groenendijk, Jeroen and Stokhof, Martin. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers*. Ph.D. thesis, University of Amsterdam, Department of Philosophy.

Halpern, Joseph Y., Moses, Yoram, and Vardi, Moshe Y. 1994. "Algorithmic Knowledge." In *Proceedings of the 5th Conference on Theoretical Aspects of Reasoning and Knowledge (TARK'94)*, 255–266. Morgan Kaufmann.

Hamblin, Charles Leonard. 1973. "Questions in Montague English." *Foundations of Language* 10:41–53.

Hintikka, Jaakko. 1962. *Knowledge and Belief: An Introduction to the Logic of Two Notions.* Ithaca: Cornell University Press.

Hoek, Daniel. 2022. "Questions in Action." *Journal of Philosophy* 1–23.

—. forthcoming. "Minimal Rationality and the Web of Questions." In Dirk Kindermann, Peter van Elswyk, and Andy Egan (eds.), *Unstructured Content*, 1–30. Oxford: Oxford University Press.

Lewis, David. 1979. "Attitudes *De Dicto* and *De Se*." *Philosophical Review* 513–543.

Lewis, David K. 1988. "Statements Partly About Observation." *Philosophical Papers* 17:1–31.

Parikh, Rohit. 1987. "Knowledge and the Problem of Logical Omniscience." In Z. W. Ras and M. Zemankova (eds.), *International Symposium on Methodologies of Intelligent Systems*, 432–439. North-Holland.

Pérez Carballo, Alejandro. 2016. "Structuring Logical Space." *Philosophy and Phenomenological Research* 92:460–491.

Soysal, Zeynep. 2022. "A Metalinguistic and Computational Approach to the Problem of Mathematical Omniscience." *Philosophy and Phenomenological Research* 1–20. doi:10.1111/phpr.12864.

Stalnaker, Robert C. 1984. *Inquiry.* Cambridge, MA: MIT Press.

—. 1991. "The Problem of Logical Omniscience, I." *Synthese* 89.

Teague, Richard. forthcoming. "The Problem of Closure and Questioning Attitudes." *Synthese* 1–24.

Williamson, Timothy. 2009. "Probability and Danger." *The Amherst Lecture in Philosophy* 4:1–35.

Yalcin, Seth. 2018. "Belief as Question-Sensitive." *Philosophy and Phenomenological Research* 97:23–47.

—. 2021. "Fragmented but Rational." In Cristina Borgoni, Dirk Kindermann, and Andrea Onofri (eds.), *The Fragmented Mind*, chapter 6, 156–179. Oxford: Oxford University Press.