Conceptual Engineering And The Access Problem for Metaphysical Possibility

Sharon Berry

October 3, 2025

Outline of talk

Table of Contents

Access Worries re: Metaphysical Possibility

- **Access worry for knowledge of metaphysical possibility** How could we know truths about metaphysical possibility given we only observe the actual? (c.f. Kant)
  - Wouldn't human accuracy about metaphysical possibility require some kind of spooky coincidence?
  - c.f. similar worries about knowledge of mathematics or morals (as traditionally/realistically construed)

Agenda

In this talk I'll:

- Review and raise a boundary-setting problem for Williamson's answer.
- Offer a moderately deflationary alternative that combines:
  - Realism about certain **ingredient facts** (logical possibility, physical/nomic possibility, conceptual joint-carvingness, physical facts about the actual world) relevant to metaphysical possibility.
  - **Mild conventionalism** about how to go from these ingredient facts to verdicts about the space of metaphysically possible scenarios.

My favoured solution to access worries will combine:

- Chalmers-inspired 2D structure to explain how empirical discoveries can matter to metaphysical possibility
- Realism about many ingredient facts relevant to logical possibility judgments
- Mild conventionalism about the function (from ingredient facts to beliefs about the space of metaphysically possible scenarios)
- A practical function story about why we might still care about metaphysical possibility facts which contain this conventional element.

Note

This is exploratory work combining elements from different projects. My goals today:

- press a specific problem for Williamson,
- sketch an alternative framework
- get your reactions.

Table of Contents

Background re: Access Problems

I take it that a realist philosophy of domain D **faces an access problem** to the extent that adopting this view [1] seemingly forces one to posit

- some mysterious 'extra' match between human beliefs about D and certain belief-independent facts,
- which intuitively cries out for explanation but goes unexplained.
- proponents of less realist view of domain D can avoid positing

So an access problem is a kind of 'how possibly' question: how could one possibly explain the match between human psychology and abstract mathematical reality?

[1] together with uncontroversial scientific and philosophical facts

Answer By Toy Model

We can address the access problem (like other 'how possibly' questions) by providing a toy model, i.e., a sample explanation of human accuracy about D facts as the realist understands them which

- is intuitively satisfying/coincidence banishing
  - i.e. removes appearances that accepting realism about D commits one to positing **extra** spooky coincidence – which could be avoided by taking an alternative approach to D
- includes all the features of our situation which seem to make adequate explanation of human accuracy about realist mathematics impossible
- (but may be simplified in other ways)

Note: the point that answering access worries only requires dispelling impressions that accepting realist knowledge of D requires positing *extra* spooky coincidences.

For example, when answering mathematical access worries...

We can't just explaining that we reliably get true mathematical beliefs by accepting true mathematical axioms and use of FOL (a conditionally reliable process) to derive further beliefs

- For even if true, this just pushes the bump in the rug – explaining away one apparent spooky extra coincidence mathematical realists must posit (true mathematical beliefs) by appeal to another (true mathematical axioms).

But we can appeal to accurate sensory and memory faculties that all parties to the debate are committed to positing

- Empiricists like Mill and Quine can takes our possession of working sensory faculties for granted
  - (they don't have to simultaneously answer 'veil of perception' access worries about knowledge of the external world, or goodman access worries about abduction)
- Formalists can take working faculties of memory and syntactic manipulation for granted (access problems about these would be access problems for everyone)

Williamson on Knowledge of Metaphysical Possibility

So let's start with Williamson's proposed answer to access worries [**?**]. He suggests

- Practical/evolutionary pressure to get *nearby* counterfactuals right, by simulting offline
  - (e.g. If I were to throw the pebble, it would land in the lake')
- If we can assess counterfactuals generally, then we can get metaphysical possibility knowledge via

$$\Diamond\varphi \text{ iff } \neg(\varphi \mathrel{\Box\!\!\rightarrow} \bot).$$

  - Why? if $\Diamond\phi$ is false then all (zero) closest possible worlds in which ϕ are ones where $\bot$. If $\Diamond\phi$ is true then clearly $\neg(\varphi \mathrel{\Box\!\!\rightarrow} \bot)$
- So facility with counterfactuals yields facility with metaphysical possibility claims.

Closeness vs. Boundaries

Worry: Selection for realistic hypotheticals fixes a *closeness ordering* among nearby worlds, but not the *boundary* of metaphysical possibility.

- Getting “what happens if I throw the rock” right just requires adequately caputring closeness relations among mundane possibilities.
- But metaphysical possibility judgments often hinge on which *physically necessary* laws to *bracket* when considering weird antecedents.
  - *Bracketable:* e.g. laws of motion / geometry .
  - *Non-bracketable:* e.g. color-space structure, analyticities.

So it seems questionable that Williamson's selection pressure could explain our ability to draw the boundaries of metaphysical possibility correctly

- (if we take these boundaries to reflect a reference magnetic joint in nature, as presumably he does).

Table of Contents

Early Modern Approaches to Epistemology of Metaphysical Possibility

Similarly, classic early modern remarks on the epistemology of metaphysical possibility aren't very helpful with access worries.

In the *Enquiry* Hume (in effect) suggests we can recognize the metaphysical possibility of the sun not rising tomorrow from the fact that this scenario

- can be distinctly "conceived" (mentally pictured/imagined?)
  - That the sun will not rise tomorrow is "is conceived by the mind with the same facility and distinctness, as if ever so conformable to reality."
- doesn't "imply... contradiction"
  - "That the sun will not rise to-morrow is no less intelligible a proposition, and **implies no more contradiction than the affirmation**, that it will rise. We should in vain, therefore, attempt to demonstrate its falsehood.

Appeals to Mental Picturing

But attempts to explain knowledge of metaphysical possibility

- via facts about conceivability qua something like psychological facts about what we can mental picture + opinions about acceptable Wittgensteinian 'methods of projection'
  - (i.e., such that entertaining a mental picture which depicts p under these methods of picturing counts as conceiving of p and motivates belief P is metaphysically possible)
- just pushes metaphysical access worries back to a question about how we recognize appropriate methods of projection (given psychological facts about we can mentally picture).

Appeals to Not Implying a Contradiction

And appeals to whether a hypothesis 'implies contradiction' face a worry similar to the one for Williamson above.

- Getting knowledge of MP in this way would require determining which actual laws of nature are bracketable
  - braketable: laws of motion and the structure of physical space
  - not braketable: laws of the structure of color space and analyticities
- And we face an equally serious access worry about how we do that. So this just pushes the bump in the rug (apparent need to posit *some* spooky extra coincidence remains).

To highlight the substantiveness of this remaining access problem, note

A. Some paradigmatically metaphysically impossible claims like ‘something is crimson but not red’ don’t **logically imply** contradictions as straightforwardly formalized.

- Some logical positivists hoped to uncover deep structure true analysis which would make these coincide.
- But how do we recognize when we’ve got the right deep structure, except by relating this to prior knowledge of logical possibility? [2]

[2] An alternative might be to adopt an element of conventionalism, as I will advocate.

B. We can sometimes fruitfully and interestingly reason from metaphysically impossible hypothesis in a way that avoids (straightforward) contradiction and explosion

- Arguably forcing arguments in set theory provide a disciplined, mathematically interesting, classical logic friendly way of exploring metaphysically impossible scenarios where facts about second order logic (and thereby the continuum hypothesis) are different
- Medieval logic exercises where you were given a metaphysically impossible hypothesis and had to stave off direct contradiction for as long as possible (?)

So noting connections to mental picturing and learning what 'implies contradiction' doesn't solve access worries about knowledge of metaphysical possibility.

Table of Contents

A First Thought About Avoiding the Problem for Williamson

Stage 1 (Problem): Williamson explains nearby counterfactuals, not *boundaries*.

My Fix: Add mild conventionalism about boundaries.

Upshot: Dissolves “which laws to suspend?” but creates *two new worries*.

Two new worries created by conventionalism

Q1 (Discovery): How can empirical discoveries (water = H_2O, geometry) matter if boundaries are partly conventional?

Q2 (Point): If MP is semi-conventional, why have it at all?

Roadmap

Re: Q1 Make room for relevance of science via Chalmersian 2D Generalized: facts about the actual world helps determine the *space* of metaphysically possible scenarios, (not just the exstension of terms like ‘water’ at metaphysically possible worlds).

Re: Q2 Suggest that our metaphysical possibility talk usefully helps us coordinate on when to coordinate, in a way that doesn’t require it to track a deeply natural distinctions

- support this by giving examples of language games that serve the same purpose but draw different boundaries.

Table of Contents

Part 1: Making Room For Scientific Discovery

Let's start with the task of making room for empirical discovery to influence beliefs about metaphysical possibility.

In this section I'll argue that we can address this challenge in a very powerful way, while still leaving room for an element of convention.

2D Semantics 'All the Way Down'

Overview of Idea:

- Facts about the actual world don't just determine the extension of twinearthable terms like 'water' within a fixed space of worlds.
- Instead, they determine the very space of metaphysical possibilities itself.

Example of what this allows:

- Empirical discoveries (about geometry, about the constitution of water) update the "basis language" we use for recombination.
- "If the actual world had turned out to be Euclidean, our concept of MP would likely have treated all possible worlds as Euclidean".

A key idea in classic Chalmersian two dimensionalism (which makes some room for scientific discovery in learning about metaphysical possibility) is this

- Our concept of WATER doesn't just map each metaphysically possible world to the set of things that count as water in that world.
  - e.g. It's not just a function from each possible world to the set of H_2O chunks in that world.
- Rather our concept WATER fixes a function from each centered way the world could be w, to an extension for 'water' if w is the actual world. This function takes
  - centred possible worlds where the watery stuff near the center is H_2O to a function that maps each possible world to the set of chunks of H_2O.
  - centred possible worlds where the watery stuff near the center is XYZ to a function that maps each possible world to the set of chunks of XYZ in that world.

I'm suggesting a similar idea, but (of course) not taking facts about metaphysically possible worlds for granted.

The idea is that just as our concept WATER includes the idea that

- if the watery stuff around us is H_2O then WATER is coextenstive with H_2O at all metaphysically possible situations
- if the watery stuff around us is XYZ then WATER is coextensive with XYZ

Our concept of METAPHYSICALLY POSSIBLE SITUATION might involve something like

- if all the elegant principles of geometry humans are a priori attracted to hold with physical necessity, then
  - the range metaphysically possible structures for space = range the logically coherent hypotheses about space which satisfy all these principles.
- if only a certain subset of these principles are actually physical laws, the range of metaphysically possible worlds is broader,
  - all metaphysically possible scenarios must satisfy this subset
  - but they can violate other such a priori attractive principles (e.g. Euclid's axioms read as claims about paths in physical space)

Ingredient facts

So one might think of our conception of metaphysical possibility as determining a function from some **ingredient facts** including the following

- the state of the actual world
- objective physical/nomic possibility
- logical possibility
- naturalness joint carvingness of concepts (practices/extensions)

to verdicts about the space of metaphysically possible scenarios.

I take access to these other ingredient facts[3] to be sufficiently popular/independently motivated to be presumable when considering whether knowledge of metaphysical possibility creates a special access problem.

[3] Except possibly for logical possibility facts, see appendix for details on the answer to access worries about those I propose in prior work.

Thus we can see answer access worries about knowledge of logical possibility via appeal to a combination of two sources:

- Realist Ingredient Facts: Knowledge of the actual world, physical possibility, logical possibility, and conceptual naturalness.
- Conventional Transition Function: A (semi-)conventional choice of how to go from ingredient facts to a choice of a "basis language and recombination principles" which determine out a space of metaphysically possible scenarios (described in the basis language)

Access Worry Answering Strategy & Payoff

So we might get the following **Division of Labour in Answering Access Worries:**

- Accuracy About Ingredient Facts: We might ultimately owe substantive (IBE, correction by experience, etc.) accounts of access to the *ingredient facts*
  - but giving this answer is a problem for everyone/something we independently need.
- Accuracy about the 2D function: Access worries concerning correct ingredients-to-metaphysical possibility boundaries *transition* are answered **metasemantically**.
  - Just as 'fuschia' picks out whichever shade our practice coherently applies it to, 'metaphysically possible' picks out whichever modal notion our transition-practice associates with the way ingredient facts are in the actual world.

Payoff

Thus combining realism about ingredient facts with enhanced Chalmerisan 2 dimensionalism lets us combine

- a deflationary metasemantic answer to access worries about boundary setting above (how do we know which physically necessary laws can be bracketed?)
- significant realism about metaphysical possibility judgments
  - e.g. accommodating intuitions that empirical work, not arbitrary convention fixing, is required to determine whether 'water is possibly H_2O' is true.
  - it would be possible to be undetectably wrong about metaphysical possibility via being undetectably wrong about chemistry, as above.

One more detail: Williamson style counterfactual judgments may still play a role

...in generating metaphysical possibility knowledge, but a much less direct one.

- **Williamson's picture**:

  Counterfactual judgements: "if ϕ were true then $\perp$"
  $\rightarrow$ metaphysical possibility claims

- **My picture**:

  Ingredient fact beliefs
  $\rightarrow$ *conventional boundary* (choice of basis language + recombination principles) $\rightarrow$ judgements characterising the space of metaphysically possible scenarios via the basis language
  + Counterfactual judgments "if [basis language description] were true then ϕ"
  $\rightarrow$ general metaphysical possibility beliefs (using vocabulary outside the basis language)

Table of Contents

Why would we care about metaphysical possibility if it's semi-conventional?

Now let's turn to the second challenge:

- if that's all there is to metaphysical possibility (i.e. if it has this deflationary aspect), why would we care about it?

I will propose an answer to that question in this section.

The Practical Engineering Problem

Why Metaphysical Possibility?: It's useful to coordinate how our terms would apply in future cases – even ones we currently don't think will actually arise (e.g. if you borrowed a knife from someone who then went crazy, would justice require returning the knife?).

But there's an opportunity cost to spending time doing this coordinating (arguing/explaining how you think 'justice' would apply in those cases)

Metaphysical Possibility Practices As a Solution

Our metaphysical possibility practices often set a boundary to **which descriptions of weird scenarios to bother coordinating our reactions to** e.g,

- we wonder about how terms like ‘justice’ or ‘nation’ would apply in scenarios we judge metaphysically possible
  - but not ones we take to be metaphysically impossible e.g. ones where someone has fewer than two fingers but also more than 5 fingers, or where there is a round square table
- we seek necessary and sufficient conditions for terms applying that have no **metaphysically** (vs. logically or physically) possible counterexamples

Q: Why not just use our concept of physical possibility for this job? Spending time coordinating how to describe scenarios we take to be physically impossible seems pointless and wasteful.
A:

- But sometimes we are wrong about what's physically possible.
- So it's useful if we err on the side of caution and try to clarify/coordinate reactions to a range of cases broader than what we currently regard as physical possibilities
  - when trying to learn from reporters and scouts who encountered things we wrongly believed were physically impossible.
  - when trying to clearly and communally reason about what we should expect if alternative laws of physics held (an important thing for comparing our current physical beliefs to others and improving them)

.

Question/objection

Q: why have a **single** shared notion of metaphysical possibility/scenarios worth coordinating on?

A: Doing this lets us

- benefit from knowing others are seeking to clarify and coordinate usage in the same range of cases, e.g., get reason for confidence from other people not worrying that something has been left ambiguous
- Share the quest for **(metaphysically) necessary and sufficient conditions** for the application of certain terms.

When coordinating on when to bother coordinating, we'd plausibly want to select a range of descriptions with the following three features.

1. (near) certain to include everything practically/physical possible
   - despite the fact that humans very commonly form temporary false opinions about what's physically possible.
2. particularly simple to evaluate
   - (so we don't lose more time on coordinating when to coordinate, than we save focusing our efforts at coordination)
3. rules out as many things as physically impossible things possible given 1, 2

I'll suggest that our metaphysical possibility practices seem nicely fitting (if not optimal) for satisfying these three goals.

- Example: arguably **recombination intuitions** about metaphysical possibility, reflect a convenient solution to the practical engineering problem above

Here are some examples of what I mean by recombination intuitions about metaphysical possibility.

- Tractatus: metaphysically possible situations exactly correspond to truthtable rows specifying which tractarian facts do/don't obtain
- Post tractatus intuition that there are (not necessarily distinct) metaphysically possible scenarios 3D euclidean space corresponding to every combinatorially possible way of choosing
  - for each location in the 3D space whether or not that location is occupied by a point particle (binary choice)
  - for each point particle, a positive real number for the mass of that particle etc.

Lewissian: any way of stitching together intrinsic duplicates of two metaphysically possible objects/regions yields something metaphysically possible

Our having concept of metaphysical possibility partly anchored by recombination practices like those above can be seen as

- a solution to the practical engineering problem of sweeping out a range of descriptions which is both
  - in the goldilocks zone above (broad enough to almost certainly include the actual world, but narrow enough to not waste too much time contemplating the bizzare)
  - but **easy to apply**

  when 'coordinating on when to coordinate' as described above.

However, things get very complicated (partly reflection how our actual ways of evaluating metaphysical possibility are quite complicated)

- So I'll now try to flesh out the above picture by describing how something like our practices extends a series of simpler language games which
  - share the same 'coordinating on when to coordinate' function
  - but pick out different modal notions than metaphysical possibility

Table of Contents

What's the Use? Engineering a Coordination Tool

- **Function:** Metaphysical possibility serves as a tool to focus **communal attention**.
  - It enables **epistemic self-correction** by exploring beyond current physics.
  - It helps **social coordination** by stress-testing how we apply terms.
- **The Problem:** We can't waste time debating every bizarre, logically possible scenario.
- **The Solution:** We need a **"Goldilocks" boundary** for which scenarios to take seriously—broad enough for safety, narrow enough for efficiency.

To see how such a Goldilocks boundary can be engineered –and make the above considerations more concrete let's consider a range of simpler language games, leading up to something like our actual metaphysical possibility practice.

Game 1: Pluralism in Practice (Business AIs)

- **The Setup:** Imagine **disembodied AIs** negotiating contracts. Their success depends on shared interpretations of terms (e.g., what qualifies as "a stakeholder"? "need to know?") for planning and reputation. Human intuitions about space or mental picturing are irrelevant to them
- **Their Method:** They define "legit" scenarios by:
  - Choosing a **basis vocabulary** of core terms (‘pays‘, ‘promises‘, ‘discloses‘).
  - Agreeing on **analyticities** connecting terms in the basis language.
  - **Recombination based picture**: Any logically coherent description of a scenario in the basis that satisfies the analyticities is a case worth coordinating on

Moral 1 (Pluralism)

The coordination function is served even if their boundary is very different from ours. **It’s the role that matters, not the specific rules.**

Game 2: Human-Specific Needs (Footsoldiers)

- **Initial Setup:** Coordinate on direct **sensory inputs** (e.g., what counts as "war elephants"?) to align fast reactions. This is quick but brittle.
- **Improvement:** To enable better planning, the basis shifts to stable **objects and properties** (distributions of color, temperature, etc.) inferred from sensory input.

Moral 2 (Human Specificity)

The specific, messy details of our practice are shaped by our **embodied needs and cognitive architecture** (senses, memory, planning).

The General Recipe: Our Recombination Intuitions

- **Moral 3 (Generality):** These games reveal a general recipe:
  1. Pick a **basis language** (business terms, sensory properties).
  2. Allow all logically consistent **recombinations** within that basis.
- **Our Practice:** Philosophical "recombination intuitions" (a la Lewis, Tractatus) are just our sophisticated version of this same recipe. It's a cognitively cheap way to generate a Goldilocks space.

The Final Twist: A Posteriori Updates

- Getting to our actual metaphysical possibility practice requires adding (at least) one more special feature: we **update our basis language** (that in turn determines the space of hypotheses we consider worthy of coordinating on) based on what we learn about the actual world.
- **Example:** Learning "heat *is* molecular motion" (/ physically necessarily coextensive with it) changes the basis we use for recombination. We no longer treat "heat" and "motion" as fully separable elements.

This leads directly to the Chalmersian 2D model advocated above...

...where facts about the actual world help determine the very space of metaphysical possibilities we coordinate on.

So overall, I think this story lets us answer the two worries for pure deflationism about metaphysical possibility I began with.

- Scientific discovery objection: "But claims like water is (metaphysically necessarily) H_2O, heat is (metaphysically necessarily) molecular motion seem like deep facts about nature, not conventions"
- Philosophical Significance objection: "But seeking necessary and sufficient conditions for knowledge, justice, etc. is substantive philosophy, not just convention-mongering"

I take my proposed answer to the first challenge, but let me briefly recap/clarify how I want to answer the second challenge.

Answer to Philosophical Significance Objection

Even if the boundaries of metaphysical possibility are partly conventional, debates about necessary/sufficient conditions can be substantive because they're really debates about:

- What follows from the ingredient facts (logical possibility, conceptual naturalness, physical laws) we've coordinated on using given our conventions
- How our terms apply across the range of scenarios we've agreed to care about and try to coordinate our usage regarding.
- Whether proposed analyses respect the conventional structure we've adopted

Analogy with chess:

Although we've conventionally adopted certain rules (chosen to have a social practice of chess rather than chess*), questions like "Can a queen move diagonally?" or "is there any position which guarantees mate in 245 moves but no fewer?" have objective answers.

Similarly, given that we favor a certain view of how ingredient facts determine the space of metaphysically possible scenarios (via a choice of basis language and recombination), questions like "Is knowledge necessarily justified true belief?" have non-arbitrary answers reflecting how our terms apply across the space of scenarios we've agreed to care about.

Table of Contents

Conclusion

In this talk I have

- Criticized Williamson's answer to access worries about knowledge of metaphysical possibility
- Suggested an alternative, mildly deflationary Chalmersian 2D account of metaphysical possibility, and answer to access worries about our knowledge of it
- Proposed an account of the usefulness (in conceptual engineering terms) of our notion of metaphysical possibility, compatible with this deflationary

Comparison with Sider

In [**?**] Ted Sider also advocates (something like) reducing/analysing facts about metaphysical possibility to certain ingredient facts (with resultant gains in theoretical parsimony and reduction of access worries).

- I agree with him in not taking metaphysical possibility to be fundamental
- But I disagree with his efforts to eliminate *all modal facts* from the fundamentalia.
- Specifically, I'd suggest taking facts about logical possibility as ingredient facts is helpful because
  - logical possibility intuitions seems to key role in metaphysical possibility judgments via humean recombination intuitions
  - an answer to access worries re: logical possibility is arguably independently needed[4], and substantive answers have been proposed (c.f. [**?**][**?**]).

[4] to account for mathematicians' ability to recognize logically coherent posits, in answering mathematical access worries

Takeaways

- Metaphysical possibility earns its keep by guiding attention and coordination of language use.
- Its boundary is *engineered*: anchored in actuality/physical possibility, narrower than logical possibility, cognitively workable.
- Access worries are reduced/deflated: metaphysical possibility facts systematically reflect more tractable ingredient facts + (semi-conventional) construction rules.

Appendix: Access to Logical Possibility I

In [**?**] I propose to account of our acquisition of powerful correct methods of reasoning about logical possibility via processes like

- **Actual → Possible**: What we observe to be actual is thereby known to be logically possible
- **Abduction/IBE**: Patterns in what's actual can support inferences about what's logically necessary/impossible (e.g., never observing 9 distinct combinations of 3 binary properties suggests logical impossibility, not just physical law)
- **Generalization/Reflective Equilibrium**: Wrong general principles adopted via something like abduction can be corrected when found to conflict with knowledge of what's actual and

Appendix: Access to Logical Possibility II

Some key points:

- This story uses faculties mathematical skeptics already accept (observation, abduction)
- I'm not suggesting this process is infallible, but neither is observation,
  - and mathematicians' confidence does actually seem to decrease from paradigmatic mathematical certainty when considering axioms about higher set theory (as my abductive story predicts).

See [**?**] for details about recognizing the possibility of infinite structures, evaluating the logical coherence of (in effect) second order conceptions of mathematical structures