Kaplan's Paradox and Epistemically Possible Worlds

David Chalmers

1. Epistemically possible worlds

Metaphysically possible worlds: S is metaphysically possible iff S is true in some metaphysically possible world.

Epistemically possible worlds: S is epistemically possible iff S is true in some epistemically possible world (or: in some scenario).

Here: S is epistemically necessary iff S is a priori. S is epistemically possible iff ~S is not a priori.

Epistemically possible worlds (scenarios) can be understood as centered metaphysically possible worlds, given 2D assumptions about the plenitude of such worlds. Alternatively, they can be understood as a modal space in their own right.

E.g. a scenario = an equivalence class of epistemically complete sentences (epistemically possible sentences S such that there is no sentence T such that S&T and S&~T are both epistemically possible) in a canonical language. A specification of a scenario is a sentence in its equivalence class.

A sentence S is true at (or: is verified by) a scenario w iff D⊃S is epistemically necessary, where D is a specification of w. We can define a relation ver(w, s) between scenarios and sentences that takes on the same range of truth-values as sentences. Then the (epistemic or primary) intension of S is a function from scenarios to truth-values, mapping w to ver(w,s).

One can argue that scenarios and intensions satisfy the plenitude principle above as well as other principles such as compositionality, and that intensions so-defined have many nice properties: e.g. the quasi-Fregean property that 'A ≡ B' is a priori iff A and B have the same intension. Useful for Fregean sense, narrow content, attitude ascriptions, epistemic logic, subjective probability, indicative conditionals.

2. Kaplan's Paradox

Kaplan's paradox: the following plausible claims form an inconsistent triad.

(i) there are at least as many propositions as sets of worlds

(ii) there are at least as many worlds as propositions

(iii) there are more sets of worlds than worlds

The case for (i): there's a proposition true at all and only those worlds.
The case for (ii): there's a world where that proposition is uniquely asserted.
The case for (iii): Cantor's theorem.

The same goes for epistemically possible scenarios:

(i) there are at least as many intensions as sets of scenarios

(ii) there are at least as many scenarios as intensions

(iii) there are more sets of scenarios than scenarios

3. Responses

Lewis's response: some propositions are not asserted (or entertained) at any world. Asserting/entertaining a proposition requires a functional role, and there are no more than beth-3 functional roles.

But: if we allow infinitary minds that can entertain infinite conjunctions and disjunctions, it's easy to generate more than beth-3 functional roles. Assume we have κ states corresponding to distinct thoughts. Then one can generate 2^{κ} functional roles for conjunctive thoughts and another 2^{κ} for disjunctive thoughts. For any of these thoughts p one can generate a new functional role for the thought that someone is entertaining p. So Lewis needs to disallow arbitrary infinite minds – on what grounds? Some might appeal to a brute necessity. But such minds seem epistemically possible, so that route is not open on the current approach.

Kaplan's response: ramify the spaces of propositions and worlds. Level-0 propositions are about extensional matters, level-1 propositions are about these and level-0 propositions, and so on. Level-n worlds are maximal level-n propositions.

But: this rules out a uniform semantic treatment for natural language sentences: For any level n, "John asserted a level-n proposition" doesn't express a level-n proposition. Also, it is arguable that all possible level-n propositions are necessitated by possible level-0 propositions (cf. physicalism), in which case there are no more level-n worlds than level-0 worlds.

Regarding both responses: one can raise related worries without mentioning propositions. E.g. arguably: for all κ it is (epistemically) possible that there are κ atomic entities with binary-valued properties. This yields 2^κ worlds. Or if qualitatively identical worlds are excluded, even noting that it is (epistemically) possible that there are exactly κ atomic entities will generate more worlds than any cardinal. This line of reasoning suggests that for a cardinal κ, there are more than κ worlds, so that there can be no set of all worlds. The moral of this situation and Kaplan's paradox is the same: there are too many worlds to form a set. But the Lewis and Kaplan responses give no purchase on the more general worry about cardinality.

I take the upshot of both Kaplan's paradox and this simple cardinality puzzle to be that we should deny (iii). The scenarios (and the worlds) are analogous to the sets: they are in some sense indefinitely extensible, and cannot be collected into a set. Then just as there are not more sets of sets than sets, there need be no more sets of worlds than worlds. Cantor's theorem does not apply when the entities in question do not form a set.

But then: what does the space of scenarios look like? And how can we understand intensions if there is no set of all scenarios?

4. A stratified construction of scenarios

I'll construct a stratified picture of the scenarios, with different spaces of scenarios for every infinite cardinal κ.

Assume an infinitary language L with a countable lexicon (consisting at least of a basis for epistemic space, and possibly arbitrary invariant expressions in natural language), and infinitary constructions that at least include arbitrary infinite conjunctions and disjunctions (and perhaps arbitrary infinite sequences of quantifiers). I'll take it that sentences of L correspond to thoughts that could be entertained by infinitary beings, and are assessable for epistemic possibility and necessity.

For any infinite cardinal κ, a *κ-sentence* is a sentence of length less than κ. Then there will be at most $f(\kappa)$ κ-sentences, where $f(\kappa)$ is the sum of ω^α for all cardinalities $\alpha<\kappa$. (If the Generalized Continuum Hypothesis is true, $f(\kappa)=\kappa$ for all κ.) A *κ-conjunction* is a conjunction of at most $f(\kappa)$ κ-sentences. A *κ-complete* sentence is an epistemically possible κ-conjunction d such that for all κ-sentences s, $d\&s$ and $d\&\neg s$ are not both epistemically possible. A *κ-scenario* is an equivalence class of κ-complete sentences, each of which is then a specification of that scenario.

E.g. the ω-scenarios, where ω is the cardinality of the integers. An ω-sentence is a finite sentence. There are ω (= $f(\omega)$) ω-sentences. An ω-conjunction is a conjunction of at most a countably infinite number of finite sentences. An ω-scenario will be an equivalence class of ω-complete conjunctions of this sort.

κ-Plenitude says: If S is epistemically possible, S is verified by some κ-scenario. κ-Plenitude is false if the sentences include μ-sentences for $\mu > \kappa$, but is arguably true when restricted to natural language sentences uttered by finite beings. This claim follows from the following principles

(E1*) If s (in natural language) is epistemically possible, s is implied by some κ-conjunction in L.

(E2*) If a κ-conjunction s (in L) is epistemically possible, s is implied by some κ-complete sentence in L.

These claims, especially (E2*), are nontrivial, but one can argue for them, and prove them from reasonable assumptions. [The start of such an argument for (E2*): Say that s is κ-completable if it is implied by some κ-complete sentence in L. If S is true, S is κ-completable as S is epistemically necessitated by the conjunction of all true κ-sentences in L. This reasoning is a priori. So if s is epistemically possible, it is epistemically possible that s is κ-completable. So one cannot establish a priori that s is not κ-completable. But uncompletability of S is knowable a priori if knowable at all, and unknowability can be excluded by plausible assumptions about apriority. N.B. S5 for apriority is required.]

So the set of κ-scenarios serves as a sort of epistemic space in its own right. One can define κ-intensions for all natural language sentences, with nice properties. Some κ-scenarios will correspond to scenarios (simpliciter), while some will correspond to many scenarios (simpliciter), when the relevant κ-conjunction is κ-complete but not epistemically complete. But κ-completeness is good enough for most purposes where natural language is concerned.

Illustration: consider the space R^R of functions from real numbers to real numbers, and by considering ω-sentences and ω-conjunctions characterizing such functions in arithmetical language (without terms for arbitrary reals). Many such functions (including any continuous or computable function) can be specified uniquely with a ω-conjunction. But not all: there are c^c members of R^R (where c is the cardinality of the real numbers) but only c ω-conjunctions. For any function f, we can take $D(f)$ to be the conjunction of all ω-sentences satisfied by f. Then $D(f)$ will be ω-complete. But often D(f) will be true of many other (discontinuous, incomputable) functions that cannot be distinguished from f using finite sentences. So many scenarios will correspond to equivalence classes of functions that cannot be distinguished in natural language.

To do semantics for English, it probably suffices to invoke the space of ω-scenarios, and ω-intensions defined over this space. The distinctions that this space misses are not distinctions we can express or entertain. Unlike Kaplan's construction, this construction will have no problem handling sentences about propositions, and no problem handling sentences describing arbitrarily large universes, as long as the sentences themselves are finite. And arguably an ω-conjunction can be used to uniquely specify the actual world.

If we apply Kaplan's paradox to κ-scenarios and κ-intensions, the case for (ii) is removed: for most κ-intensions, there will be no κ-scenario in which that intension is entertained. If we apply it to entities that are κ-scenarios for some κ and κ-intensions for some κ, (iii) is false, as before.

What's missing? ω-Plenitude (and κ-Plenitude generally) will fail for some infinite sentences: there will be infinite sentences that are epistemically possible, but whose ω-intension is not true at any ω-scenario. Q: Can we recover a picture with scenarios (simpliciter), intensions (simpliciter), and a plenitude thesis that applies to all possible sentences?

5. Scenarios and intensions simpliciter

A scenario (*simpliciter*) is an entity that is a complete κ-scenario for some κ. Here a complete κ-scenario is one specified by a κ-conjunction *s* that is not just κ-complete but epistemically complete.

Q: What are intensions? Not sets of scenarios: any set of scenarios is a set of κ-scenarios for some κ, so there will be no everywhere-true intension, and all intensions will be everywhere-false sufficiently high in the hierarchy. Not functions, construed as sets of ordered pairs of scenarios and truth-values, one for each scenario: there are no such sets.

One might appeal to proper classes, holding either that intensions are classes of scenarios, or classes of ordered pairs of scenarios and truth-values. If classes construed as set-like collections individuated by their members (but too large to be sets), this will not help. Cantor-style reasoning will then suggest that there are more intensions than scenarios, and it will remain epistemically possible that any intension can be entertained at a scenario, so we will regenerate Kaplan's paradox. One will then get a hierarchy of various sorts of collection -- classes, metaclasses, and so on – with a corresponding hierarchy of intensions. But then the previous situation with sets will apply to on the new situation with collections. Better to hold that all such collections are sets, and to draw the morals that one would have to draw at the level of collections with sets.

A better alternative: understand intensions as functions without identifying functions with sets of ordered pairs. This is already familiar from set theory: there are many functions defined over all sets, such as the function that maps an arbitrary set s to {s}. The axiom of replacement is often case in terms of such functions. Such functions cannot be identified with sets of ordered pairs. Rather, they correspond in some fashion to defining formulae.

On the current picture, we have a well-defined two-place function ver(w,s) that maps arbitrary scenarios and sentences to truth-value. For any given sentence s, this function delivers a well-defined one-place function mapping scenarios w to ver(w,s). Of course neither of these functions are sets of ordered pairs, but they are well-defined functions all the same. Their nature might perhaps be construed using extended versions of set theory, such as Fine's ("Class and Memberhip", *Journal of Philosophy,* 2005) theory of classes, on which the relevant entities are individuated by defining conditions rather than by members. (Also: Linnebo's (2008) related theory of properties, or an analog of Bealer's algebraic view of propositions on which they are simple entities individuated by operations they are involved in.) But what matters for our purposes is that talk of functions in this sense is coherent.

Does Kaplan's paradox rearise? One might think that there are more such functions than worlds for Cantorian reasons. But this is incorrect: the notion of function used here yields only a function for every defining condition. Even on Fine's plenitudinous construction, there are no more defining conditions than sets. So there are no more functions over possible worlds than possible worlds.

There remains Kaplan's "constructive" version of the paradox (used by Bruno Whittle ("Epistemically Possible Worlds and Propositions", *Nous*, 2009) to raise problems for epistemic space): one takes the function from propositions to worlds and diagonalizes it, yielding a proposition that is true in a world iff the unique proposition asserted in that world is false there: then in a world where this proposition is asserted, paradox results. I think that this is a version of the Liar paradox – the relevant unique assertion will be something like "The unique proposition asserted in this world is false" – and so it is a problem for everyone, not just for epistemic space.

I think everything I have said about epistemically possible scenarios also applies to metaphysically possible worlds. One can have stratified sets of κ-worlds, with corresponding κ-intensions. κ-worlds can arguably do the work of worlds for the purposes of possible-world natural-language semantics (perhaps not for possible-world metaphysics?). Worlds simpliciter are κ-worlds for some κ. There is no set of all worlds. Intensions/propositions simpliciter are (non-set-theoretic) functions from worlds to truth-values. There are no more such functions than worlds, so Kaplan's paradox is defeated.