

The Table of Sampling Assumptions in Anthropics

[Sandberg](#): “.... is like anthropic reasoning: complicated, error-prone, looks a bit magical but once you start getting real evidence much of it becomes irrelevant.”

The meta-meta-problem of anthropics is how to find the right way to choose between principles of selecting the correct sampling assumptions, but before we get there, let's first list all known and even unknown sampling assumptions.

The full list, including some related ideas, has 38 item (google doc [here](#), text version below).

I also suggest several possibly new assumptions for the sake of completeness of the list, but it doesn't mean that I endorse them (like “intensity sampling” and “observer-moment-size sampling”) and it doesn't mean that they were never suggested before.

<i>The main principle of a group of assumptions:</i>	<i>Short abbreviation and link:</i>	<i>Full name and definition:</i>	<i>Problems and arguments:</i>	<i>Doomsday argument-like problems:</i>	<i>Sleeping Beauty solution:</i>	<i>Proponents:</i>
1. The role of possible vs. real observers.	SSA	<i>Self-sampling assumption.</i> I am randomly selected from all <i>actually</i> existing observers of my reference class.	The problem with choosing the reference class; Many other sampling assumptions are just SSA for a specific class; Different reference classes have different types of the end.	Classical doomsday argument: Civilization will end soon (but it is not problematic unless we have too good expectations about the future). Adam and Eve: probability pumping via manipulating the number of the future observers (but maybe evolution already used this?)	1/2	
	SSSA	<i>Strong self-sampling assumption.</i> The same as SSA, but the selection is between the observer-moments (OM).		It gives too much weight to Boltzmann Brains and freak dreams, as they generate OM without proper observers.		Bostrom

	SSSSA	<p><i>Super-strong self-sampling assumption:</i> I am selected from all minds proportional to their intelligence.</p> <p><i>Variant:</i> I am selected from all observer-moments proportional to their “size”.</p>	Explains why I am not an ant.	<p>Superintelligence is impossible, or I will be it.</p> <p>Very large OMs are rare.</p>		Toby Pereira
	Universal SSA	<p>SSA is true for every reference class.</p> <p>I am randomly selected from each and every reference class to which I belong if they are mutually independent and all their members consist of qualified observers.</p>		There is a “bad corner” in the space of all possible minds where all types of the Doomsday Argument are true.		
	SIA	<p>“We should give higher probability to hypotheses that contain more observers, other things being equal”, Tomasik.</p> <p>(Classical definition: “I am randomly selected from all possible observers”.)</p>	If the universe is ergodically infinite, SIA is uninformative, as all possible observers do exist in it. Therefore, the fact, that I exist, does not give me any new information.	<p>SIA-doomsday: The Great Filter is ahead (Grace).</p> <p>Presumptuous philosopher: the hypothesis which postulates a larger number of observers could be validated without experiments.</p>	1/3	K. Grace
	SSA+SIA	<p>“But sometimes the term “SIA” also refers to “SSA+SIA”, which favors more copies of <i>you</i> specifically, without needing to define observers in general”, Tomasik.</p>	SIA tells why I am in the long (tail) branch in Sleeping beauty, and SSA tells where exactly I am in that branch.		1/3	
	Untypical SIA	<p>My uniqueness is an argument that many attempts to create me were made. E.g. many stars exist and many people.</p> <p>It is similar to SIA, but it is now not an assumption, but a theorem, which could be proved.</p> <p>It works only if the given actually existing observer is “untypical”, and thus its creation requires more attempts.</p>		<p>It is an argument for Rare Earth: a larger number of planets which tried but failed to create life – is more probable than a smaller number of such planets.</p> <p>Multiverse is real and it is as an explanation of the fine-tuning.</p>		<p>A.Turchin</p> <p>I invented it to illustrate SIA</p>

2. Selection only in some tempo-spatial aspects.	Now-time selection principle	<p>I am randomly selected from all observers, who live <i>now</i>, but not in the past or in the future.</p> <p>A variant: I am selected from “now and the past” (depends on the model of time, like the <i>block model of time</i>).</p>		<p>Guth’s youngness paradox: I am more likely to live in the youngest bubble universe in the world with cosmological inflation.</p> <p>Therefore, no aliens exist in my observable universe.</p> <p>It also means that I should be the <i>earliest observer in our civilization</i> who learns something about anthropics (which seems to be false).</p>		
	PSA	<p><i>Physics sampling assumption</i> by Tomasik: how often given observation will appear in different universes.</p> <p>I am likely to be in the universes which have a higher <i>density</i> of observers like me.</p> <p>Variant: The same, but also taking into account the size of the universes (the regions with the highest density could be relatively small)</p>		<p>PSA favors <i>Big Crunch</i> over <i>Heat Death</i> as it has a smaller total volume.</p> <p>PSA favors <i>solipsism</i> as it gives the highest density of observers.</p> <p>In MWI, most of my copies appear just before Big Rip.</p> <p>Panspermia is almost proved, as it ensure higher spatial density of civilizations.</p>		Tomasik
	MWI-selection	<p>I am randomly selected from all MWI branches, counting each branch separately (with observers).</p>		<p>Anti-youngness paradox: I am now in the last moment of the universe’s existence (e.g. just before False Vacuum decay) as the number of branches is maximal at that moment. Quantum immortality seems to balance the problem, as I will constantly miraculously survive the end of the universe.</p> <p>Alternatively, if we count not the number of the branches, but the total measure of all branches as constant, then the measure declines extremely quickly for any given branch and I must be very early.</p>		

	Entropy-based selection	The worlds with high complexity are thermodynamically improbable and rare.		We are the most the complex civilization ever, decline will follow. Technological progress will experience unexpected obstacles aimed at lowering the complexity of civilization. Most Boltzmann brains are simple.		Strugatsky in “Definitely maybe”.
	Civilizations-sampling assumption	We are a typical civilization . (Hanson in the Grabby Aliens).		Grabby aliens are at the distance of around 1 Gyr.		
3. The role of the observers' self-reflection in sampling	Qualified observers	I am randomly selected from the set of <i>qualified observers</i> : the ones who are thinking about anthropics. Variants: <i>Wider</i> : Or at least <i>can</i> think. <i>Thinner</i> : Or are thinking <i>exactly</i> like me.		The end is <i>very soon</i> , as the qualified observers appeared only recently (around 1970) in our civilization's history and the number of them is growing.		
	Minimal reference class	Minimal reference class by Bostrom: subjectively indistinguishable agents.				
	A question defines the answer.	Question-defined reasoning: “The room is green because I am asking why it is green”.		After a right question, the world disappears “ The Nine Billion Names of God ”		
4. Decision-theoretic approaches	FDI	<i>Functional decision theory</i> : I am randomly selected from all functionally similar lines of thought. (This is my interpretation of FDT, could be wrong – AT)	Counterfactual power over what others do. Example: If I decide to vote for candidate X, others will also do it.	Gives too much weight to non-human computations; However, as I am human, no such computations in stones, computers or AIs.		Yudkowsky
	ADI	<i>Anthropic decision theory</i> (a variant of Updateless Decision Theory). I should reason in a way which increases global utility, without updating on my local position.	I should ignore my copies in simulation and BB as they have zero impact, even if they are numerically dominating	Utility monsters? I should reason as if I am a future creator of FAI: in that case, I would have the biggest possible impact.		

	FNC	<p><i>Full non-indexical conditioning</i></p> <p>Chances of my appearance under different theories are:</p> <p>"$P(\text{theory of the universe} \mid \text{I exist}) \propto P(\text{I exist} \mid \text{theory of the universe}) * P(\text{theory of the universe}).$"</p>				
5. Computational approach	KSA	<p>Computations sampling assumption – similar to UDASSA, based on Kolmogorov complexity to weight measure.</p>				
	UDASSA	<p>OMs are selected proportionally to their Kolmogorov complexity.</p>		"Simpler" minds are more often.		
	"Law without law" article	<p>In this theory "I" and "others" have different statistics.</p> <p>In this theory, observer states are connected via computational complexity of the transition from one to another in the math universe.</p> <p>It is similar to UDASSA, but the difference is that UDASSA calculates the probability of each OM, and LwL computes the probability of transition from one OM to another.</p>		<p>Simulations are unlikely: a person who is often observable is not the one who is often experienced from the inside.</p>		

6. The difference between subjective and objective probabilities	Perspective-based reasoning	Self-sampling is impossible as it confuses the first person and third-person perspective. “Because self-locating probability has no underlying process, assumptions treating the first person as the sampling outcome are needed to fill this gap.”		I am special, not typical.		dadadarren
	Anthropic trilemma approach	Subjective probabilities are completely different from objective and could be manipulated by the order of creating of copies.		Anthropic devil : a measure monster which manipulates subjective probabilities in unobservable ways.		
7. Other similar principles	Copernican mediocrity principle	“In physical cosmology, the Copernican principle states that humans, on the Earth or in the Solar System, are not privileged observers of the universe”				
	Bland indifference principle	“bland indifference principle... prescribes indifference only between hypotheses about which observer you are, when you have no information about which of these observers you are...”				Bostrom
	Unification of theory of identity (Tomasic)	I am represented by all my copies simultaneously, therefore, no selection happens and no need for the selection assumption.	No problems like “I am BB” or “I am in a simulation”.	Impossibility to predict the single unitary future: any future outcome has probability 1. Bad outcomes dominate.		

8. "Measure" as a probability of being a given observer.	The number of microstates	Each macro state of mind (OM) could be represented by many microstates. The <i>measure</i> is proportional to the number of microstates which represent one macro state (similar to entropy).		Minds with the largest number of microstates should have the highest measure. These are minds with large sensory fields, but simple interpretation states. We are them, so there are no minds with even larger sensory fields. Also "meditating" minds with large but "empty" sensory fields are more likely??		
	Qualia-selection	I am randomly selected only from those who: a) have qualia b) have "my type" of qualia (my red) c) has mental qualia are about anthropics		Humans are Qualia monsters: we have the most intense qualia. Fine-tuning to generate mental qualia is no more needed: The universe existed only for me to think this exact thought and may disappear after that. Qualia world: only qualia exist, no other objects.		
	Causal connectedness. Selection should be only from a causally-connected region of the universe	I am randomly selected only from observers in the same causally-connected region of the universe, but not from other universes.	Other universes are just possible ones.	I am in the universe with the largest causal connection between observers (like cosmological inflation)??		

	Energy-based selection	I am selected based on the energy used for computations; I am unlikely to be a Boltzmann brain, as they “use” minimal energy.		<p>Very energy-inefficient, “thick” brains will dominate. Thus simulations are less probable, but “zoos” are still likely.</p> <p>I am more likely to be a biological brain than a simulation.</p> <p>“Measure monsters” and anthropic devils hunt on observers in the multiverse by increasing the energy spent on observers’ copies.</p>		
	Intensity-based selection	My experience is in the middle of the set of all minds ordered by the <i>intensity</i> of pain and pleasure.		<p>S-risks and paradise engineering (wire-heading, hedonium) are very rare.</p> <p>Life is generally boring.</p>		
9. Related ideas	Observations selection	Not observers are selected, but observations.		Anthropic principle and fine-tuning.		
	Universes’ sampling	Replicating (via black holes) fecund universes from evo-devo theory . Self-replicating universes are the most numerous ones.		We will self-destroy ourselves via creating a black hole which will give rise to many universes with the same fine-tuning.		
	Panpsychism	All physical processes and computations have qualia and I am randomly selected from all of them.		<p>Anti-panpsychism: there should be some cut-off in the complexity below which experience is not possible, and it is near the human level. But if we add the “qualified observers” requirement, it becomes this cutoff.</p> <p>Waterfall argument: I am Boltzmann brain.</p>		

	<p>Gaia hypothesis (Quantum immortality-in the-past favors effective survivors)</p>	<p>The idea is similar to the Gaia hypothesis, which claims that the Earth has a special ability to survive climate instability. In the same way, if I survived many rounds of quantum suicide, I likely have some special survival ability.</p> <p>Stronger version: I am more likely to be an observer who <i>will</i> eventually get a “measure burst” in the future.</p> <p>This assumes retrocausality of measure, may be reinforced by loops-in-time.</p>		<p>I am more likely to live in the period of time near Singularity, AGI and life extension, and to be interested in anthropics.</p> <p>Stronger version “Transcendental advantage”: my personal life story is untypical and evolves in the direction of becoming a “God”: a being of high complexity and measure.</p> <p>(This claim seems less crazy if we replace “I” with “planet”: The Earth has a special history compared to other planets and will probably evolve into superintelligence.)</p> <p>If loops-in-time similar to <i>unescapable prophecies</i> are possible, I am stuck in one of them, as it the significantly increases measure.</p>		
	Me-sampling	<p>I am randomly selected from all minds who think that they are the same observer, e.g. Alexey Turchin, but have different OMs (Identity-selection assumption).</p> <p>In some sense, it is opposite to SSSA which selects from the same OMs, which could be in different observers.</p>		<p>Eventfully, my personality becomes the most typical in the universe.</p>		
	Uncertainty about a valid sampling method	<p>We should take an average of different sampling methods proportional to our credence in them (or based on the prediction market).</p>		<p>Meta-Doomsday argument: We should multiply our credence in DA on its predictions.</p>	2/5 in SB	

	Superposition of uncertainties	Similar to above, but assumes that this uncertainty is ontological, not epistemic (similar in some sense to the Heisenberg principle).		The universe behaves in such a way that I can't choose the right sampling method, because if I can, the size of my reference class will become extremely small.		
--	--------------------------------	--	--	---	--	--

There are too many assumptions and I feel lost. How could it be solved?

On the meta-level, there are several principles how to choose between three dozen of sampling assumptions from the table:

- Choose just one selection principle and hold it as a “correct”. However, how can we prove that it is true? What is the source of the rightness of a selection principle?
 1. Test it experimentally.
 2. Find unbeatable mathematical proof.
 3. Kill all other assumptions as logically inconsistent.
- Ditch selection. Anthropics doesn't work at all.
- The difference between selection principles doesn't matter in most cases, except infinities and many copies situations. Don't be afraid to update.
- All of the sampling assumptions are valid and we are located in some very peculiar part of the universe where different types of Doomsday argument are all true.
- The right choice of assumption depends on the subtle details of the question asked.
- One should use meta-level uncertainty over different assumptions and assign them different credence.

The meta-meta problem

The meta-meta-problem of anthropics is how to find the right way to choose between principles of selecting the correct sampling assumptions.

Tomasik wrote about the assumption choice problem:

“Maintaining model uncertainty

Each anthropic view seems to have its own problems:

- SSA yields the doomsday argument, Adam & Eve, Lazy Adam, etc. (and might yield solipsism if it didn't cheat by not paying for its reference class).
- SIA yields the presumptuous philosopher.
- FNC tells us we should be in an infinite universe and then provides no guidance beyond that.
- A principle of indifference over copies of ourselves, even given modal realism, can't decide between certain hypotheses about the universe where it seems obvious we should update our beliefs based on evidence.
- PSA yields an arbitrary-seeming Presumption of Denseness and may (or may not?) imply solipsism when combined with Solomonoff probabilities.
- KCA tells us we were more likely to have been George Washington than a random, non-notable person. Or, if that's not the case, it at least arbitrarily favors people whose birth ranks are simpler numbers.”

There are two ideas in anthropics: “selection assumption” and “measure”. On a closer look, they are similar, as both of them define chances to be a given observer. Measure is a function which defines selection, therefore, by defining what is measure, we define the selection assumption. Example: In SSA, my measure is equal between each member of my reference class, and is 0 for any non-member. Measure is more general than selection assumption as it gives non-equal chances for different minds. Example 2: If there are two minds, but I my subjective chances to be in the first one 2 times higher than in the second, we could say that the measure of the first mind is 2 times higher. Some of the selection assumptions are defined via measure, for example SSSA, which says that I am more likely to be a mind with higher intelligence – so selection is happened over all actually existing minds, but with different weights.

Assumption is bad epistemic. They must be proved and empathically verified.