Pascal's Hell

Petra Kosonen

In the beginning, on a small planet in the Solar System, in the Milky Way galaxy…

Satan: I have an offer for you, Pascal, as I have heard that you might be interested in a small probability of a huge payoff.

Pascal: Anything that maximizes expected utility!

Satan: Great! And your utility function is unbounded, am I right?[1]

Pascal: Yes, and additive in terms of people's happy days of life.

Satan: Excellent. So, the offer is this: I will flip a coin, and if it lands on heads, I will help humanity settle on new planets in faraway galaxies and live in bliss until the heat death of the Universe.[2] Until the heat death happens, it will be like heaven. But if the coin lands on tails, then everyone on Earth will suffer excruciating pain for the next fifty years. That's the offer. If you decide to accept it, I will return to Earth every fifty years and give the same offer until either you (or your descendants) refuse the offer, the coin lands on heads, or the Sun expands and makes life on

---

[1]The utility function does not necessarily have to be unbounded for this case to work—it is enough that the upper bound is very high and the lower bound very low.

[2]This is the fate of the Universe in which Pascal and Satan live.

Earth impossible. If you decide not to accept the offer, humanity will live its earthly existence as mere mortals until life on this planet is no longer possible (humanity will not be able to expand out from Earth without my help!)

Pascal: Your offer sounds great—even odds of Utopia! And if we don't win this time, we'll almost certainly win eventually.

Satan: Oh, pardon me, I forgot to say that my coin is somewhat biased. If you accept all the (20 million) offers, the probability of heads happening at least once is one-in-a-googolplex. I admit the odds aren't great. But if the coin lands on heads, I will create a thousand googolplex happy Earth-like planets.

Pascal: Not to worry, the offer is still amazing. The expected value of taking those gambles is clearly greater than the expected value of rejecting them. Actually, its expected value might even be greater than the expected value of the offer I initially thought you were making… So, I'm positively surprised.

Satan: Oops, I made a mistake. I read the wrong page. The instruction manual (*Creating Hell*) says that the probability of heads ever happening on Earth is one-in-Graham's-number. But it is in my power to create any finite number of happy Earth-like planets, so I believe I can still give you a good offer. If the coin lands on heads, I will create a million Graham's number of happy Earth-like planets.[3]

Pascal: Now your offer is even better! Although I dread the almost certain torture for everyone on Earth for the next billion years, the expected value of your offer is far greater than the expected value of not taking it. So, rationality compels me to accept it.

---

[3]The Universe Pascal and Satan live in is much larger than our Universe.

Pascal and Satan then agree on the deal, and Satan flips the coin. Unsurprisingly, it lands on tails.

Satan: You and everyone on Earth will now suffer excruciating pain for the next fifty years.

Pascal: Oh well. I made the right choice, given the information I had. And the future is still great in expectation. Thank you for your offer.

Satan: I'm always happy to help. See you again in fifty years!

Pascal: See you in fifty (long) years! You are always welcome here.

Satan: I never imagined persuading people to enter (finite) hell would be this easy…

* * *

So Satan traveled from one planet to another, and the inhabitants of those planets—also expected utility maximizers with unbounded utilities—always accepted his offer. And they all lived happily ever after (in expectation). But according to Satan's instruction manual, the probability of the coin ever landing on heads was merely one-in-a-googolplex, so the Universe was almost certain to be void of joy and laughter.[4]

---

[4]This dialogue is based on Pascal's Mugging by Bostrom (2009), which in turn is based on informal discussions by various people, including Yudkowsky (2007). Pascal's Mugging is similar to Pascal's Wager, except that the former does not involve infinite utilities. Pascal (1958) famously argued that one should believe in God because of the possibility of gaining an infinitely good payoff in Heaven: "Let us weigh the gain and the loss in wagering that God is. Let us estimate these two

References

Bostrom, N. (2009), 'Pascal's Mugging', *Analysis* **69**(3), 443–445.

Pascal, B. (1958), *Pascal's Pensées*, E. P. Dutton & Co., New York.
  URL: *https://www.gutenberg.org/files/18269/18269-0.txt*

Yudkowsky, E. (2007), 'Pascal's Mugging: Tiny probabilities of vast utilities'.
  URL: *http://www.overcomingbias.com/2007/10/pascals-mugging.html*

---

chances. If you gain, you gain all; if you lose, you lose nothing. Wager, then, without hesitation that He is."