

What is philosophy?

Peter G. Doyle

Version dated 6 June 1997
GNU FDL*

In the early 90's I went to my first and only philosophy lecture. This was a special Distinguished Lecture, in celebration of the debut of Princeton's special program in moral philosophy. The speaker was Robert Nozick, from Harvard, and the topic was 'the alien banker's paradox'. This paradox, which apparently is quite well known, usually goes under the name of its inventor, who despite having a name like that of a nineteenth century mathematician turns out to have been some modern guy working someplace like JPL. Anyway, since I don't remember the name, I'm going to call it the 'alien banker's paradox'.

Here's how it goes. An alien comes down and amazes you by repeatedly and reliably predicting what you're going to do next. Maybe he writes down what you're going to do, and after you do your thing you look at what he wrote, and by God he was right. Once he has convinced you of his powers, he gives you two boxes. One box is transparent and you can see that it contains \$100; the other box is opaque. Now the alien gives you a choice: You can keep the contents of either the opaque box alone, or of both boxes. The alien informs you that if you choose the opaque box alone, you will find that it contains \$1,000,000; but if you choose both boxes, the opaque box will turn out to be empty, and you will only get the \$100 from the transparent box. The reason you're supposed to regard this as a paradox is that you're supposed to figure that the million already either is or is not in that opaque box, and that being so, you might as well take both boxes.

*Copyright (C) 1997, 2006 Peter G. Doyle. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, as published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

So what do you do? Well, if you're Robert Nozick, what you do is you develop a very general theory that involves analyzing the situation in three different ways, and then taking some kind of numerical combination of the three different arguments. Or something like that: Actually, I didn't understand much of what he said. About all I can distinctly remember from the rest of his lecture is that during the question period he more than once declined to address some topic or other while 'standing on one foot'. Since the lectern was blocking my view, I couldn't tell whether he was actually standing on one foot, but I tend to doubt it—I think this was just a figure of speech.

I attended this lecture in the company of my friend, colleague, and guru John Conway, the colorful mathematician best known as the inventor of 'surreal numbers' and 'Conway's Game of Life'. During the question period, Conway observed that 'hard cases make bad law', but Nozick either did not agree, or felt that this saying did not apply in the present case.

Now as I tried to follow Nozick's lecture, I was naturally trying to think for myself about this paradox. Since I knew no philosophy, all I had to work with was common sense, and the kinds of arguments I came up with are such as would occur to any sensible person. For instance, if you really believe this alien knows his stuff, then you're going to get a million, so go with it. Sure, maybe you could get an extra hundred, but that would make the alien look like a fool, and your first priority here should be not to anger this alien. I remember a TV commercial from the 60's which dramatized the dangers of fooling Mother Nature into mistaking margarine for butter. How could the extra hundred offset the danger of angering the alien banker?

And then suddenly, about half-way through the question period, it hit me. I tore two sheets of paper from my pad and folded them into makeshift envelopes, to represent the two boxes. I took out my wallet, and extracted a one dollar bill and a twenty dollar bill. I placed the one dollar bill in the 'transparent' envelope, and either did or did not place the twenty dollar bill in the 'opaque' envelope. Turning to Conway, I let him observe the dollar in the 'transparent' envelope, and gave him the alien banker's choice. Without an instant's hesitation, Conway chose to take both envelopes. He extracted the dollar from the 'transparent' envelope and stuffed it in his pocket. If don't remember now whether he even bothered to look in the 'opaque' envelope: He knew bloody well it would be empty.