

Ockham's Razor

Issue date: 02 Sept 2007

Physicist Anthony Garrett explains the meaning of Ockham's Razor.

Anthony Garrett: The village of Ockham is about 25 miles south-west of central London, in England. It is set off the main road from London to Portsmouth, the naval town on the south coast, and nowadays the intersection of this road and the London orbital motorway lies in Ockham Wood. In this small village, seven centuries ago, a boy named William was brought up. As William of Ockham, he became a famous theologian, and today his name is known worldwide for the principle of logic called Ockham's Razor, which states, broadly, that you should prefer the simplest explanation to fit the facts. He is commemorated in his own village by a window in the parish church. I will tell William's story and the story of his razor principle, which has recently found expression in my own field, probability theory.

In his youth William joined the Franciscan order, founded by St Francis of Assisi less than a century before. The Franciscans were a breath of fresh air in the mediaeval church, bringing a new emphasis on the beauty of the created world and choosing personal material poverty. They were known as the grey friars, from the colour of their clothing. William, after training in London, went on to Oxford, which was then a great Franciscan centre of learning. He became the leader of a radical school of thought called Nominalism. The theology of the time seems excessively pedantic today; a later parody of it is the debate over how many angels can dance on the head of a pin. But the so-called scholastic theologians were the intellectual pioneers of their day, seeking to construct a unified moral, legal, theological and political framework for mediaeval life. William's sharpness of mind led him to be labelled by an 18th century commentator the 'razor of the nominalists'. So the first reference to a razor, many centuries later, is to William himself.

William went on to argue that scientific enquiry should not be limited by theological directions from the church, so he joined the great debate about faith and reason. Personally I see reason as building on faith, just as a mathematical theorem begins from axioms and proceeds using reason, to its conclusion. Whatever, William's claim was too radical for a church that saw itself as embodying divine power, and he was summoned to the papal court to answer charges of heresy. William must have made an enemy, for the charges had been laid by the Chancellor of Oxford University himself. William reached Avignon in France, where the papacy had been temporarily displaced by European politics, in the year 1324. The enquiry dragged on four years until, with the Pope and his court becoming increasingly hostile to Franciscan austerity, William fled by night to join a secular opponent of the papacy. This was Louis of Bavaria, leader of the German princes. In his city of Munich, William was able to return to teaching and ignore the papal decrees against him. He died, probably of the Black Death, which took a terrifying one-third to one-half of Europe's population, about 20 years later in 1349. In William's last years, Louis' power declined and then he died, and William sensibly attempted reconciliation with the papacy. He was buried in the Franciscan church in Munich, and now the little road Occamstrasse commemorates him in a fashionable area of the city. (This spelling is from his Latin name Occamus, in contrast to the English Ockham; both versions are common today.)

So much for William's life and times; what of his razor? Well it is taken to refer to the principle of parsimony that, from the Latin, 'it is vain to do with more what can be done with less', or 'a plurality of things is not to be posited without necessity'. In other words, keep it as simple as possible. Sherlock Holmes says as much. The first irony here is that William was actually more concerned with exploring the limits of this idea. The second irony is that although William often wrote of the idea, he was certainly not its inventor. Later writers attached his name to it, but the idea was common among scholastic theologians, and goes back much further. Here is a quote from Ptolemy, writing in 2nd century Alexandria about changes in the earth's solstices and equinoxes, called Procession. 'It is a good principle to explain the phenomena by the simplest hypotheses possible, insofar as there is nothing in the observations to provide a significant objection to such a procedure'. This is exactly the idea that is now translated into probability theory, as we'll see.

The ironies continue. The idea is most famous from a Latin aphorism, 'Entia non sunt multiplicanda praeter necessitatem', broadly, ideas should not proliferate unnecessarily. When most scholars were taught Latin, until just 30 or 40 years ago, this was a commonplace phrase. Although it was generally assumed to have been written by William, it was in fact a summary from half way between his time and ours, by an Irish scholar, John Ponce. The idea was finally given the name 'Ockham's Razor' as recently as the 19th century, by William Hamilton, changing the use of 'razor' from William's mind to the principle itself, presumably because it cuts away unnecessary complication.

So much for the history of the idea. In the 20th century it has found mathematical expression within probability theory. So: is this mathematical version the final, precise version of the idea, settled for all time? No, it isn't. Mathematics is precise, and words have shades of meaning which make them ambiguous; mathematics and words are suited to different things. So today's version of the Razor in probability theory, corresponds to one interpretation of the words, but other interpretations could generate different mathematical realisations. The idea made precise in probability theory is a very important one though, relating theory and practice in some very basic areas of physics. Let me illustrate this.

The new mathematical razor applies whenever there is a set of numerical data that are polluted by noise, by processes we do not know the exact details of, and we are interested in whether there is a signal, hidden in that noise. An example that recurs as data become ever more accurate is whether there are further planets in our solar system. Look at the motion of the outermost planets such as Pluto. We know that Pluto is influenced principally by the gravitational field of the sun, round which it orbits; then by the gravitational fields of the other planets that we know about. But if you look closely enough at Pluto's motion you will still find small deviations from the motion predicted, even taking the known planets into account. The question is: do those deviations contain subtle evidence for any further unknown planets, or are they best explained as noise due to comets and meteors passing by Pluto, dust, and so on? If the answer is Yes, it is more likely that there is another planet, then where do the data suggest that this other planet is orbiting, and how big is it, what is its influence?

To see how this plugs into the idea of Ockham's Razor, imagine hunting for not just one extra planet, but two, three, four and so on. Obviously if we suppose there are enough extra planets we can fit the predicted orbit of Pluto to the observed fluctuations of the orbit very finely indeed. But the mind revolts at the idea of dozens of extra planets; it makes obvious better sense to suppose the fluctuations are due to irregular passing comets, dust, and so on. The physicist Richard Feynman once said that given enough parameters (that means planets in this case) he could fit an elephant to the curve. We call this phenomenon Overfit. On the other hand, with fewer extra planets we cannot fit the data so closely, which is obviously something that you want to do. So intuitively, there is going to be a trade-off between how well you can fit the data and how many extra planets you suppose there are. In other words, how complicated the theory is. This trade-off between goodness of fit to the observations, and how simple the theory you're using is, is made precise in the new mathematical razor. It allows us to say how probable it is that there is zero, one, two or more undetected planets; and if one or more, what is the best guess of where their orbits are and their masses. Both the number of extra planets, and their positions and masses, are chosen so as to allow the best fit to the data. Ockham's Razor is not just 'choose the simplest theory that fits the facts', but 'choose the simplest theory that fits the facts well', and there is a measurable trade-off, between goodness of fit and simplicity of the theory; a trade-off between flexibility and economy.

Ockham's Razor is also the motivation behind unification of physical theory. A good example of this came nearly 100 years ago. The German physicist Max Planck had invented an early version of the quantum theory that explained a baffling phenomenon: the speed of electrons that were thrown off when light is shone at a metal. His equations called for a new physical constant, a new constant of nature, whose value had to be found from the observations he made. But the same idea was then applied to explain the amount of radiation given off by a hot body, an electric fire, for example, and also to explain the wavelengths of light that are absorbed by hydrogen atoms. But of these further phenomena had been experimentally studied and each had required its own physical constant of nature to be set separately from the observations. The new idea related these two extra constants to Planck's and accurately gave their values. Three supposedly separate phenomena had been shown to have the same underlying explanation. The quantum idea was rapidly accepted in consequence.

My last example is from cosmology. When Einstein worked out his general theory of relativity and gravity early in the 20th century, and improved on Newton's venerable theory, there was room for an arbitrary constant, known as a parameter, in his equations. To keep things simple he was tempted to put it to zero, but another consideration weighed even more heavily: he believed on philosophical grounds that the universe was unchanging on the large scale. He believed it was unchanging in how the great clusters of stars, called galaxies, relate to each other. This meant that his number could not be zero, for technical reasons.

But some years later, it was found that the galaxies were in fact all rushing away from one another. In Einstein's mind, an informal version of the Ockham analysis immediately took place and he reverted to the value zero for his number, which is called the 'cosmological constant' today. In this spirit, a translation of the Latin Ockham's Razor, 'entia non sunt multiplicanda praeter necessitatem' would be 'Parameters should not proliferate unnecessarily'. This particular plot has thickened though: the value of Einstein's cosmological constant is once again in question. Is it zero, or is it very small, and should be chosen so as to best fit the data?

We don't know yet. This is why these questions are exciting.

It is a long way from the modest 13th century village of Ockham to modern research laboratories with state-of-the-art technology, proving the secrets of elementary particles and cosmology. Our link is William, and the principle he wrote about which allows us to improve our answers to questions about the universe, according to the data that is coming out of those laboratories. I think he would be pleased.