Heuristics: Rules of Thumb

Tony Starfield

recorded: November, 2009

What is a heuristic? A heuristic is a rule of thumb. It is something that is sometimes true and sometimes works, but sometimes it doesn't work. Heuristics are needed in situations where you don't have a clear set of rules, but you do have guidelines as to how to do something better or to accomplish something more economically or quickly.

Modeling is something of an art. If one could say when you are developing a model, "Step one, do this. Step two, do that. Step three, do the next thing," then one wouldn't need heuristics. But because it is an art, one needs to have a set of heuristics to think about and perhaps guide you when you're trying to develop a model or getting stuck in the development of a model.

The heuristics that I'm going to go through were collected over a large number of years, largely from making mistakes. They come from experience, and each time one has a bad experience or sometimes a good experience, one stops and says, "What works? What doesn't work? How could I do this better next time?" And, out of that comes a heuristic.

This is something you might want to think about in your professional life. It really helps one to evaluate everything one does and say, "What have I learned from it, and what heuristics can I pass on to people who might be doing this in the future?"

The list of heuristics that you have is not in any special order, but I'll go through it, and remember as I go through it that sometimes a heuristic is important and useful, and sometimes it just doesn't apply to the situation in hand.

The first heuristic is always useful. Keep it simple. Use the leanest model for the purpose at hand. This ties in with the idea of rapid prototyping. When you first build a model, build the simplest model that could possibly meet the objectives you have. You can always make it more complicated later.

The next heuristic, number two, is plan your output. It's useful before you build your model to say, "Suppose I develop this model, what kind of output would I want to get and how would I

present it? Would I be looking for a graph? Would I be looking for a table? What kind of numbers would I want to present?" This guides you in the way in which you develop your model.

Heuristic number three ties right in with this. Go one step further - anticipate your results. If you're planning to use a graph, sketch what you think that graph is going to look like. This is useful because, when you actually run your model, if you get something similar to your expectation, that's great. If you get something very different from your expectation, that forces you to ask yourself the question, "What was wrong with the way I was thinking about this, or alternatively, is there a bug in my model that I haven't found yet?"

Heuristic number four is sometimes useful and sometimes not. Look for upper and lower bounds. When you are trying to solve a problem, think about what the worst answer might be and what the best answer might be. Sometimes those two answers are close enough that you don't really need to go any further and develop a model. Other times the distance between them is such that you really do need a model and looking at upper and lower bounds wasn't particularly useful. Looking at upper and lower bounds when you're trying to estimate parameters or trying to guess at data is also a valuable heuristic.

Heuristic number five is something completely different. When you're thinking about a model, choose appropriate spatial and temporal scales. If you're dealing with space explicitly or even if it's implicit, if you're looking at a population that lives in an area of 20 square miles, think about whether that is the appropriate scale to be looking at. When you're thinking about time, ask whether you should be trying to project ten years into the future or a hundred years into the future. At each scale, you are going to have a different model and a different set of assumptions, because some things are important at one scale and not important at another scale.

Heuristic six says graphs, pictures and histograms are usually better than words or numbers to explain model results. People react and appreciate graphs more than they do numbers. So again, when you're thinking about how to present your data, think in terms of graphs and histograms.

Heuristic number seven is keep a list of assumptions. More than that, have the guts to make assumptions. Again, if one's doing rapid prototyping, the most important thing to do there is to simplify your model world and to keep a list of assumptions. You need to keep that list so that

you know what assumptions you have made. You need to keep that list so that you can tell other people what assumptions you have made. Much of the misuse of models comes from people not spelling out what assumptions there are, so that other people use them without realizing why they should not have used them.

Heuristic number eight is, if in doubt, leave it out. So again, if you're not sure about whether to put something into your model or not - or to say that more strongly, if you don't have a compelling reason for putting something into your model - leave it out, but add it to your list of assumptions.

Heuristic number nine is about salami tactics. What do we mean about salami tactics? The idea here goes back to when you were a kid and there was a salami lying on your mother's kitchen table. If you had gone in and said, "Can I eat that salami?" your mother would have said, "No, you're going to get sick." But if you were a smart kid, what you did was you went in and asked for a slice of the salami. And then a few minutes later, you went in and said, "That was good. Can I have another slice?" And if you kept that up long enough, you got to eat the whole salami.

How does this apply to modeling? Asking for the whole salami is equivalent to saying, "I have a problem. I am starting now, and I want to get to an answer perhaps 20 years into the future." Now trying to come up with a formula or a method to get directly from now to the answer 20 years in the future is equivalent to trying to eat the whole salami. Very often it is useful to slice the problem up. Slice it up into ten-year intervals and say, "Can I get from now to ten years? Can I then get from ten years to twenty years?" or whatever. Choose a time step, slice the problem up, and that often helps to solve the problem.

Heuristic number ten is look for useful notations. If you're developing a model, the way in which you write your equations or the way in which you represent what you're doing can often influence the way in which you do it. Some things feel right and communicate well, and other things make it harder to think about a problem.

Heuristic eleven is important. If you're going to use a formula that comes out of a textbook, maybe you go to an ecological modeling textbook and pull out the logistic equation or a Lotka-Volterra model, or maybe you go to a statistics textbook and pull out an equation. Be sure that you really understand why you are using it. Every time you take somebody else's formula or pull a formula out of a textbook and put it into your model, you are potentially bringing

in bugs. Those bugs are the assumptions behind the formula you are using. And, if you don't understand them and you haven't checked that they are appropriate assumptions for what you're doing, you're going to get into trouble.

Heuristic number twelve is about Gordian Knots. Where does the idea of Gordian Knots come from? Once upon a time, like two to three thousand years ago, there was an oracle on the coast of what is probably Lebanon today, and in that ancient oracle was a knot. And the story was that whoever could undo the knot would rule the world. Alexander the Great, when he set out to conquer the Persians, took his whole army on a detour of a couple of hundred miles specifically so, as a public relations effort, he could go to the oracle and prove that he was going to rule the world. When he saw the knot, he didn't know how to undo it, but being a quick thinker, he took out his sword and cut through it. Nobody had said how the knot was to be undone.

Well, there are many Gordian Knots in modeling. You get into situations that are really messy, and you can thrash around in those situations forever and ever, arguing this way and that way and never get out of them. When you get into those situations, think of Alexander the Great. Pull out your sword, cut through them, make a simplifying assumption, and then push ahead and come back and reconsider what that assumption meant later on after you've built the model.

Heuristic number thirteen is plan ahead for a sensitivity analysis. Every model you build needs to have a sensitivity analysis so you might as well, as you are developing the model, plan ahead for the sensitivity analysis, whether that's putting your model on a spreadsheet or just thinking about what you are going to do with your model. Bear in mind that you are going to have a sensitivity analysis.

Also, heuristic number fourteen is plan ahead for an assumption analysis. Previous heuristics have been telling you to make assumptions and have the guts to make assumptions, but at some stage you're going to have to try and consider what effect those assumptions make. So you need to think hard about how you might plan ahead when you use your model to test out your major assumptions.

Heuristic number fifteen ties in with cutting through Gordian Knots and making assumptions. It is press ahead. Don't get bogged down. You can spend a lot of time arguing about a model, but the best feedback you get is when you actually get something working and look at the results

and start performing experiments on your model. So don't get bogged down. Get a working model, but think of it as a prototype and be prepared to come back and reconsider it later and then prove it later.

Heuristic number sixteen is a sort of general heuristic. Think yourself into a problem. If you're starting with a problem straight off, and you're not quite sure where to begin modeling it, try and think yourself into it by imagining that you're the most important variable. If you're modeling a population and you're trying to figure out how to represent that population over time, think of yourself as an about-to-be-born critter in that population and imagine what your life history is.

If you're modeling a hydrological model and you're thinking of how water is going to flow over a landscape, think of yourself as a drop of water and what happens to you. If you're modeling a fire, think of yourself as a spark and where that spark might take you. This often helps you to think about how you want to model a process.

And finally, heuristic number seventeen is perhaps the most important heuristic of all - be prepared to explain your model. To build a model that you cannot explain is to lose most of the power of modeling. The reason you build your models is because the real world's complicated and you can't explain it without a model. If you develop a model that you cannot explain, you haven't gained very much.

This might seem obvious, but it isn't always obvious. Very often you come up against a problem and you don't quite know how to deal with it, and you make a mistake like pulling a formula out of a book without understanding it or putting some kind of fix into your model that gets you through a problem. If you can't explain what you're doing, you shouldn't be doing it.

< 00:13:43 END >