Toolboxes, Cards, and Individuality

Scott E Page

July 28, 2006

I need to add in here the use of Botox for stuttering, ulcers, etc...

1 Introduction

The carpenter that built our house arrived each morning in a truck that contained several cabinets. In these cabinets were the tools of his trade: a drill, a nail gun, and various sized hammers, levels, saws, and screwdrivers. As important as the tools themselves were his skills as a carpenter – his ability to use the tools. One of the themes of these lectures will be that each of us, regardless of our profession, possesses our own set of tools that we apply with varying levels of expertise. As we go through life, we acquire new tools and occasionally lose tools. Sometimes we use them correctly. Other times, we find ourselves pounding a nail with a cross cut saw.

That our personal skills differentiate us is indisputable. Some of us know how to speak French. Others of us can crochet, throw horseshoes, or mince onions. These skill sets simultaneously define us, constrain us, and guide us. They determine who are friends are, what we eat, how we play, what careers we choose, how much money we make, how "smart" people think we are, and whether we are capable of having much fun. To lead fulfilling lives, we must choose them wisely, yet leave room for caprice.

These personal toolboxes depend nature, nurture, and opportunities. We vary in our genetic ability to acquire skills. Some people can roll there r's effortlessly and speak fluent Spanish. Others, those with high concave palates, can spend hours saying "pot of tea" as fast as we can, yet never really engineer a roll. Our friends, family, and teachers also influence these tools. Growing up on a lake, not learning to swim or ice skate would be almost impossible. In some families, learning to play chess or cards is mandatory. Without even realizing it, these children learn basic combinatorics and how to perform backward induction.

Due to genetics, social pressure, or interest level we differ in our capacity to possess skills, in the rates with which we can acquire skills, in our retention of those skills, in the rates at which we can apply them, and in our abilities to use them in combination. To say someone is "smart" is to say that they have facility in one or more of these areas. Some people may be quick to learn and apply skills, and they may be have enormous skill capacities. They may speak up first in class. Others may be slow to learn and slow to apply, but they may be better at retaining. Or they may be better at using skills in combination.

Or, they may be worse at everything. Anna may be able to hold twenty skills in her head, while Bob can only hold fifteen. Anna may learn skills faster, retain more of them, and apply them more quickly. She may even be more adept at using them in combination. Does that mean she is "smarter" than Bob? In some weak sense, yes. But it does not mean that her intelligence subsumes Bob's. Bob can and probably will possess a skill set that Anna does not. He may have more friends, score higher on standardize tests, and lead a more fulfilling life than Anna. It all comes down to how their skill sets mesh with the reality within which they find themselves. Given her natural advantages, Anna is likely to outperform Bob on many dimensions, but given that the set of skill sets is so large, a certain amount of luck enters the picture.

1.1 The Deck of Cards Model

The story of Anna and Bob can be more concretely understood through simple models. Imagine that each of fifty two cards in a deck deck represents a distinct skill. Anna can choose twenty cards, and Bob can choose fifteen. We want to ask three questions of this model Question 1: how many different skill sets can Anna choose? Question 2: how many different skill sets can Bob choose? and Question 3: how likely is it that Anna skill set equals Bob's skill set plus more stuff? Or to phrase this last question more pointedly, how likely is it that anything Bob can do Anna can do better?

Anna can choose twenty cards. There are fifty cards she can choose first, fifty one she can choose second, and so on. So we just have to multiply fifty-two times fifty-one, times fifty, etc.. all the way down to thirty-three. However, we also need to take into account order if she chooses the ace of hearts and then the jack of spades, this is the same thing as choosing the jack of spades and then the ace of hearts. The number of ways the same twenty cards can be reordered equals twenty times nineteen times eighteen times . . . times one. Dividing the first number by the second, we obtain that the total number of unique skill sets Anna could acquire equals

$$\frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48 \cdot \ldots \cdot 33}{20 \cdot 19 \cdot 18 \ldots \cdot 1}$$

This equals 125,994,627,894,135, or just shy of one hundred and twenty six trillion different skill sets. As something to chew on for a minute, you might want to recall that the best estimates of the total number of people who have lived on earth are counted in billions.

If we make a similar computation for Bob, we obtain that there are a mere 4,481,381,406,320, or roughly four and a half trillion skills sets for him to choose. Finally, to answer question three, how likely is it that Anna possesses every skill that Bob possesses all we have to do is choose an arbitrary skill set for Bob and compute the number of Anna's skill sets that contain it. If we take Bob's fifteen cards out of the deck there are thirty-seven left. Anna can choose five of these. There are

$$\frac{37 \cdot 36 \cdot 35 \cdot 34 \cdot 33}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

or 435,897 ways to do this. Fixing Bob's skill set at some random collection of fifteen skills, of the 126 trillion skill sets that Anna might choose, fewer than a half million of hers contain all of Bob's skills. If we divide the 126 trillion figure by the smaller number, we get that the odds are one in about 289,046,788. There is only a one in three hundred million chance that Anna possesses every skill that Bob does.

Now of course, given that there are several billion people in the world, and if each of those people were choosing twenty cards, then the odds are that several of them would know everything Bob knows. But we should be careful here. The set of possible skills is far larger than fifty two. It is probably in the tens of thousands. Moreover, the number of skills an individual can possess is probably in the low thousands or at a minimum in the high hundreds. The deck of cards example was meant to be illustrative. The numbers of skill sets is in-comprehensively large. Carl Sagan, who himself acquired quite an impressive tool kit, would be speechless.

1.2 The Circle Model

The previous calculation of the probability that Anna knows everything that Bob knows relied on an implicit assumption on the relationship between tools. To be more precise it depended on the lack of an assumption of a relationship. In reality, tools may be represented on some sort of topology or network such as on a line or on a tree.

For example, suppose that the fifty two tools can be arranged on a line and therefore numbered from one to fifty two and that a person must acquire consecutive tools. If you know tool number seven, then in order to acquire tool number nine, you mush also acquire tool number eight. If this is true, it massively reduces the number of sets of twenty tools from over 126 trillion to thirty three. Why thirty three? Well, Anna's first tool can be number one, two, three, and so only all the way up to number thirty three. If she begins with number thirty four, then there is no way for her to have twenty tools. She could only have nineteen.

Using similar logic, Bob must have one of only thirty eight sets of fifteen tools. Interestingly, now Bob has more choices of tool sets than Anna. This is because of the severe restrictions imposed by the linear topology of tools. Now, what are the odds that Anna's skill set contains Bobs. Well if Bob's skill set contains tools number one through fifteen, then Anna's set contains Bob's only if she has the first skill in her set as well, so the odds are one in thirty three. However, if Bob begins with skill number two, then the odds are two in thirty three. If he begins with three, four or five, the odds are three, four and five over thirty three respectively. Finally, if Bob's skill set begins with skill number six through skill number thirty three, then the odds that Anna's set contains his is six over thirty three. If Bob's skills begin at numbers thirty four through thirty eight the odds that Anna's skill set contains Bob's decreases from five to four to three to two to one over thirty three. Therefore, the average probability that Anna's skills contain Bob's equals

$$\frac{1+2+3+4+5+6*28+5+4+3+2+1}{33*38} = \frac{6}{38}$$

This is approximately equal to 16%, a huge increase over the previous figure of one in three hundred million when we did not assume a topology.

1.3 Multiple Lines

The linear topology is mathematically convenient but it does not make sense in light of the hierarchical nature of many tools. If I learn how to make crusts, I can then learn how to make pies or quiches fillings. Learning quiche fillings will not help me make pies and learning to make pie fillings is of little use if I want to make a quiche. Or to consider a more technical example, if we learn how to take derivatives, we can then learn how to perform integration or how to solve differential equations. Yet, to solve differential equations, we do not need to know how to perform integration and to perform integration we do not need to know how to solve differential equations.

So, let's assume that the fifty two tools are now in two linear hierarchies. Within each hierarchy, the skills are numbered from one to twenty six. To master skill number five on either hierarchy, you have to first learn skills numbered one through four.

We can compute the odds that Ana knows all that Bob knows as follows: Bob's fifteen skills are partitioned across the two hierarchies. If Anna's skill set contains Bob's she still has five leftover skills to allocate across the two hierarchies. There are $N+1$ ways to allocate the N skills across the two hierarchies. Therefore, of the 21 ways that Anna can allocate her skills, for each of Bob's allocations exactly six of these contain Bob's skills. Therefore, the probability that Ana's knowledge contains Bob's is six over twenty one, or about 28%.

Now if we increase the number of hierarchies from two to three or four these numbers fall significantly, but this gets relatively complicated mathematically.

1.4 Trees

Finally, suppose that each hierarchy is a tree and not a line. Suppose that there are two trees. We won't bother with the math but it becomes far less likely than the linear case that Anna knows everything that Bob knows. Intuition as to why this is true goes as follows: Suppose that Bob allocates all fifteen tools to one tree and choose a particular branch as far as he can go. At each level, he can choose a branch so there are 2^15 or over 32,000 branches he can choose on that tree. With her six more tools, Anna is going to be able to cover at most six of those with her additional tools.

2 What do we do with this?

This all seems rather abstract. We have people choosing skill sets and we're computing whether someone is smarter than someone else. To make sense of this, first we have to distinguish between skills that require some time to master and simple tricks. Simple tricks like the rule of 72 can be learned in a second, skills like knowing linear programming or cost accounting take a while to master. So our model makes more sense in the context of skills that cannot be figured out instantly.

If we think of people as collections of skills, and, as we will learn in a few lectures, if diverse collections of skills are better collectively than homogeneous ones, then we want diverse collections of skills. But this begs the question, how do we measure diverse collections of skills? It also begs the homework question:

Homework: *Should UM give preferential admission to students with endogenous (acquired) skill diversity? Explain the incentive problems created.*