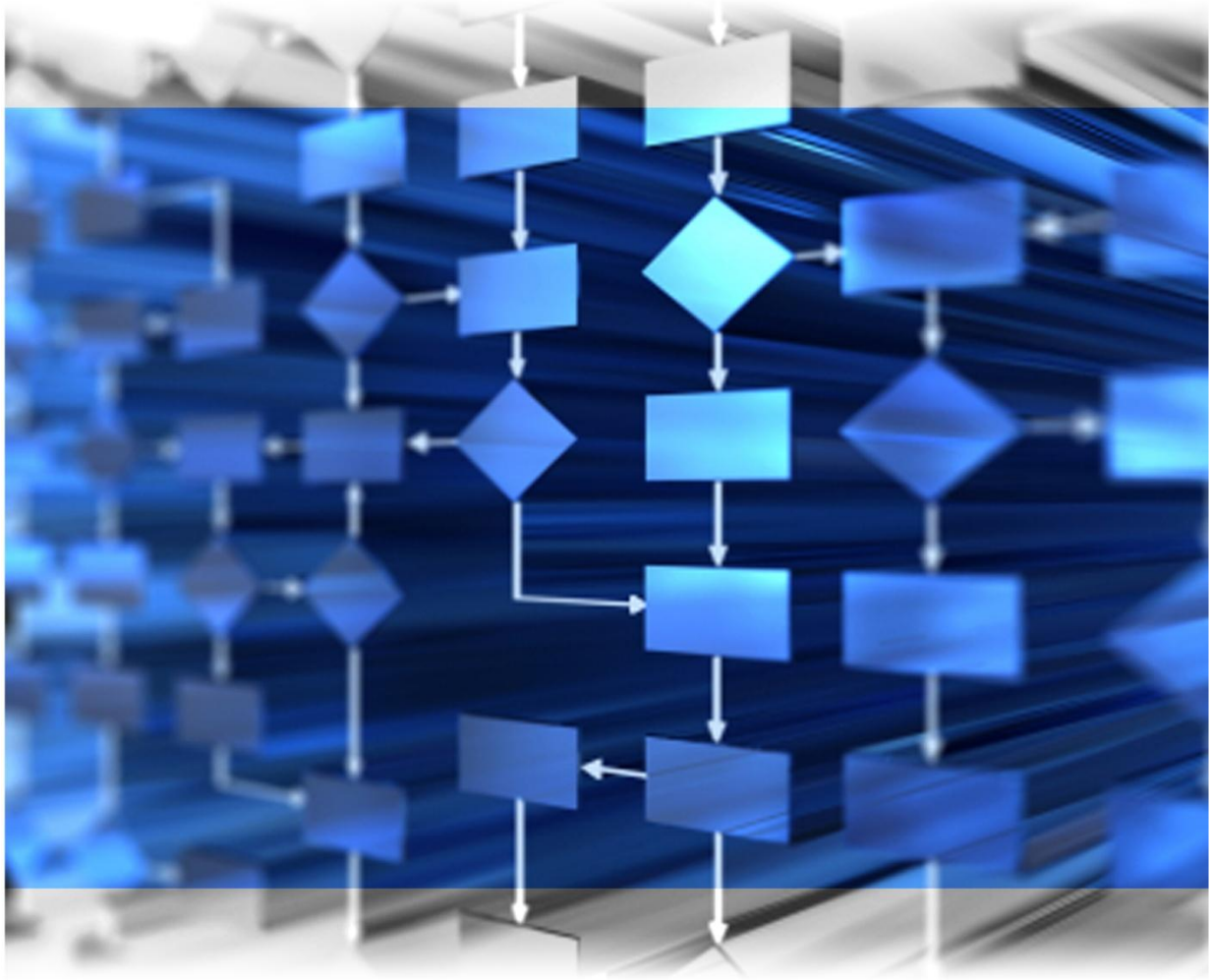
LONE STAR

THE RIGHT ANSWER



Modeling the Randomness of Humans

April 1, 2012

Modeling the Randomness of Humans

Summary

Estimating human performance and behavior is frequently required in modeling. Econometric behavior, learning, productivity, error rates and other human acts are often important factors to modelers.

Accommodating the randomness of human performance and behavior can seem to be an insurmountable challenge to building enterprise simulations, econometric models, benefit-cost analyses and other simulations. However, integrating the variability of human performance is required to create accurate models in economics, learning, productivity, error rate estimation and decision-making.

Lone Star has overcome the challenge of modeling human-based sources of uncertainty by understanding and integrating the real-world factors that affect performance variation. In Lone Star's experience, human sources of uncertainty have never posed a serious challenge to creating accurate models. This paper addresses how human randomness is successfully incorporated in TruNavigator™, StraTable™ and IRIS™ models.

Using the principles learned from Human Randomness, Lone Star has developed a model that predicts performance of human organizations with surprising accuracy.

Background

The 20th century saw several changes in the prevailing wisdom about human randomness. In the first part of the 20th century, the randomness of human behavior was nearly impenetrable. A number of scientific advances in mid-century seemed to promise that a nearly complete explanation of human variation would eventually emerge. However, by late in the century, it became clear that a deterministic explanation of human variation was probably impossible; much like the Heisenberg uncertainty principle dictates the limits of our ability to observe the behavior of atomic particles.

In 1944, John von Neumann and Oskar Morganstern published a "Theory of Games and Economic Behavior." Game Theory became a foundation for understanding strategy in quantitative, mathematical terms, breaking free from gambling examples, and moving to complex economic problems. Pioneering computer scientists, seeking mathematical explanations for human behavior, saw Game Theory as a way to explain variations and randomness as "rational" in the value framework of individual actors.

But people were shown to NOT behave in "rational" ways no matter how complex the Game Theory model. At this point, it became more widely accepted

that predicting human behavior precisely (deterministically) was probably not possible.

Late in the century, randomness in humans became less mysterious through detailed study of human behavior in real world activities. Today, we understand enough about human variability to represent humans and human organizations in models and simulations.

Today's Opportunities

One source of understanding humans is sports including the emerging field of sabermetrics. Late 20th century professional sports employed statisticians and inspired a number of amateur number crunchers. Some statistics quoted in serious tones by sports announcers have little meaning. For example, an announcer might say, "Did you know Barowski has never hit a left handed pitcher on the road when his team was behind?" This is supposed to prove Barowski has no chance of a hit, since he has traveled to face a leftie and his team is behind. But this narrow set of conditions may only represent four or five previous at bats. If Barowski were a .200 hitter, we would expect there would be several sets of batting situations where he had not yet had a hit.

If Barowski *does* have a hit under these conditions, we should not be surprised to hear the announcer tell us that Barowski is improving, that he's "hot" today or some other cliché. Sadly, as much as we enjoy the idea of "streaks", several studies have shown that hot and cold streaks are usually nothing more than the predicable random variation of a given performance level.

If we made a roulette wheel with a thousand slots, 800 for "no hit" and 200 slots for "hit," we could simulate Barowski' s batting. We would find a few cases where we "hit" many times in a row ("Barowski is hot") and streaks of many no hits ("Barowski is cold, I wonder if they'll trade him").

Of course, the wheel cannot be hot or cold; it is just providing the kind of distribution of outcomes we would get from any other similar source of random trials.

Today we know (even if the announcer doesn't) there is a great deal of information in Barowski's batting average. When we combine his information with the statistics of his teammates, we can make reasonable estimates of the team's scoring ability. We may not be able to predict when Barowski will get a hit, but we can do a respectable job of predicting the performance of his team over several games.

Learning curves are another example of human behavior. The idea of the "learning curve" is simple. People get better with repetition. This improvement, in the real world is not just "learning" and might better be called "improvement curves." The rate of improvement lessens over time, and at some point, we see little change. The concept is attributed to Hermann Ebbinghaus (1885), Arthur Bills (1934), T.P. Wright (1936), and work by the Boston Consulting Group (1960's).

Anyone who has watched a highly skilled manual assembler has been amazed at the speed of the worker. Manual shirt folding is an example. It looks so simple, but unless you have folded several thousand shirts, you won't be able to imitate the skilled folder.

The basic math behind the improvement curve is also simple. The percentage change is constant as experience grows. If a new worker assembles her first tricycle in 10 minutes, she might assemble her second in 8 minutes. If we time her fourth tricycle, we would expect it to take a little over 6 minutes.

Every time her experience doubles, she takes about 20% less time. The number of repetitions to double experience grows quickly. After 16 tricycles, she should learn to assemble the trike in approximately 4 minutes. She will need to assemble 128 of the toys to reduce the time to about 2 minutes.

We find randomness in improvement curves and among individuals, too. Dick, Jane, and Spot will not improve at the same rate. Spot may learn to fetch faster than Dick. Jane may improve in math faster than Dick. We also see variation for each individual. None of them improves exactly on the smooth curve. The performance curve theory is only true on average. In the real world, we expect to see variation (good days and bad days) among individuals, in the rate of improvement for each individual, and in other measures.

We also expect to see changes in the curve. If Sally is given new tools, say a GPS set; she may suddenly become better at fetching than Spot. If Sally has tutoring before this semester, she may show more early progress. But Dick may eventually have better math scores. Sally's head start may obscure the comparison of their aptitude (individual rate of improvement) at first. If we instruct both of them long enough, the better Math student should emerge. So, related experience and improved tools obscures the native performance and the rate of improvement of the individual or group.

Some of this randomness can be “noise” in our measurement systems. Measuring human performance is often imprecise, and measuring learning achievement is an example of an area where the randomness of the student being tested and the noise in our measurement systems can easily confound us.

Two thinkers who help us understand the nature and causes of human randomness are H.A. Simon and George Miller.

Claude Shannon’s 1948 Bell Labs paper, “A Mathematical Theory of Communication” influenced Miller and others. Miller tried Shannon’s concept of a “limited channel” on humans. This let Miller consolidate the work of other researchers. *He showed human performance is remarkably predictable when we are modeled as limited information channels*.

Miller’s work helped explain the limits of human perception. Essentially, Miller showed humans are roughly limited to about 3 bits of bandwidth for most one-dimension judgments (how long, how hot, how loud, what tone…). This relates to about seven shades of distinction. This is why most well designed surveys don’t ask for more than 7 degrees of judgment.

Miller also showed that humans integrate multiple types of input (information) to create more shades of distinction, but each additional input provides a smaller additional channel capacity. This means we have limits in terms of our ability to integrate information. This degrades each kind of information and limits how many things we can incorporate at once. However, integration does expand our perceptive abilities.

Finally, Miller showed humans have a limited capacity for immediate memory, the number of things we can recall, like the seven digits in most telephone numbers (before area codes). This applies to learning; Miller suggested we learn to lump things together. We think of each group (lump) as a single thing to extend our limits, either by explicit mnemonic methods or simply by growing familiarity. Miller’s work was extended by Atkision, Shiffin, Baddeley, Hitch, and others, but “7” remains a persistently important number describing the bounds of human performance.

The limits Miller found are important when we consider the work of Herman A. Simon. He was a polymath who, like Von Neumann, worked across several disciplines and universities. His work in the 1950’s showed the limits of human rationality and explored the limits of human cognition and attention.

Others found humans do *not* respond to Game Theory situations as rational players. In other words, we often do not play a game with the best strategy for winning. Simon offered the idea of attention economics, which helps explain the non-rational choice; humans only have so much useful attention (bandwidth), so we allocate attention to a few considerations and ignore everything else.

Taken together, Simon and Miller help explain (partly) why humans aren't rational; we ignore some input signals – sometimes we ignore most of them. More recent research shows just how bad humans are at some kinds of thinking. Probabilistic thinking (the announcer considering Barowski's batting) is better performed by pigeons than humans. Pigeons, presumably free from burdens like attention economics and cognitive bias, are more empirical than humans. In other words, humans tend to fit information into the patterns we've established, while pigeons more accurately modify their choices based on historical fact.

In the past 20 years, cognitive scientists have observed information being processed in our brains in real time. They find our "wetware" is programmed differently from one human to another (more randomness), but overall processing is remarkable similar. We respond to pleasure and pain, biasing our choice of inputs and creating a probability we may make choices in ways which prove less than rational.

Our brains are powerful pattern recognizers, but we may also see patterns that do not exist, once we begin to suspect something is true. Once we begin to "see" a pattern, our bandwidth limitations, pleasure/pain biases, and attention economics lead us to ignore inputs outside the pattern. This is why people see different things and will argue with passion, when presented the same information.

What we learn from all this is encouraging for modeling and simulation purposes:

1. Human performance is variable and not always rational, but the variation is usually well bounded.
2. Human performance change and improvement is also variable, but also usually well bounded.
3. A given human, over time and repetition is predictable in terms of the distribution of outcomes, even if we can't predict any single outcome.
4. A group of humans is *even more predictable* as their individual variation averages out, creating a group variance smaller than the average or median individual.
5. If there are conditions limiting the group's behavior and performance (corporate policy, resource limitations, training syllabi…) the behavior of

the group becomes still more predictable, even if the norms are informal group habits and traditions.

Human variability is reasonably well behaved, and the attributes of the variation (of an individual, among individuals, improvement due to change…) are usually observable.

Conclusion

Humans are indeed variable in their performance, performance improvement, and decision making processes. But the variability is limited, and the variability of groups is even more limited. Modeling and simulation of the actions and choices of people is well within the capacity of well-designed models.

Modeling human behavior in corporate, government, and economic settings is usually achievable. Since humans are bandwidth limited, the data describing their span of behavior is also limited. Impressive models can be constructed with limited data, or with data of limited fidelity. This is made easier when we model groups, and easier still when we model organizations with norms and rules.

If we decompose Barowski's batting into observable behaviors, and decompose his mental model of batting, we can determine if the batting coach changes Barowski and his teammates. We can know whether the changes are correlated with improved batting averages, runs scored, or games won.

We still can't predict whether Barowski will get a hit the next time at bat. We can't promise whether Sally or Spot will return the next fetch. These are deterministic. We *can* predict how often Sally will win, how often Barowski will get a hit, and whether Barowksi's batting coach is actually improving him.

Using these principles, Lone Star has predicted performance of human organizations with surprising accuracy. In practice, we've found the messy randomness of human institutions can be modeled with as much fidelity as most other modeling domains. We can do this because our modeling and simulation tools, TruNavigator, StraTable, and IRIS accommodate human variability.

About Lone Star

Headquartered in Dallas, Texas, Lone Star provides business and technical analysis and advisory services addressing clients' most complex, mission critical challenges. Using processes based on best practices and proprietary tools, Lone Star delivers effective outcomes and the value clients expect. Lone Star's unique knowledge and analytical capabilities and a client service-oriented philosophy means actionable results are always delivered. Lone Star's roots lie in the development, fielding and support of complex technologies and programs for the Department of Defense and

leading enterprises. For more information the Lone Star home page is www.lone-star.com.

Lone Star ● 4555 Excel Parkway ● Suite 500 ● Addison, TX 75001

Tel: 972-690-9494 Fax: 972-690-9495 www.lone-star.com