

Mathematical Model Reveals the Patterns of How Innovations Arise

The work could lead to a new approach to the study of what is possible, and how it follows from what already exists.

by Emerging Technology from the arXiv January 13, 2017

Innovation is one of the driving forces in our world. The constant creation of new ideas and their transformation into technologies and products forms a powerful cornerstone for 21st century society. Indeed, many universities and institutes, along with regions such as Silicon Valley, cultivate this process.

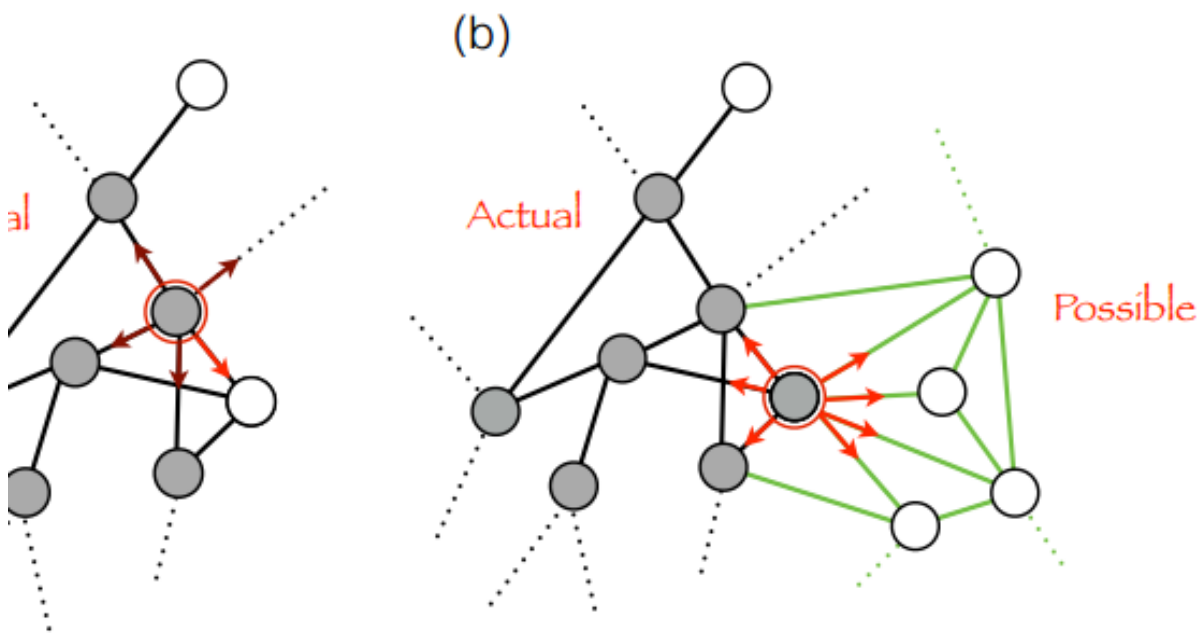
Recommended for You

- 01 Questionable “Young Blood” Transfusions Offered in U.S. as Anti-Aging Remedy
- 02 A Woman in Nevada Died from an Unstoppable Superbug
- 03 A 100-Drone Swarm, Dropped from Jets, Plans Its Own Moves
- 04 Poker Is the Latest Game to Fold Against Artificial Intelligence
- 05 One Man's Quest to Hack His Own Genes

And yet the process of innovation is something of a mystery. A wide range of researchers have studied it, ranging from economists and anthropologists to evolutionary biologists and engineers. Their goal is to understand how innovation happens and the factors that drive it so that they can optimize conditions for future innovation.

This approach has had limited success, however. The rate at which innovations appear and disappear has been carefully measured. It follows a set of well-characterized patterns that scientists observe in many different

circumstances. And yet, nobody has been able to explain how this pattern arises or why it governs innovation.



Today, all that changes thanks to the work of Vittorio Loreto at Sapienza University of Rome in Italy and a few pals, who have created the first mathematical model that accurately reproduces the patterns that innovations follow. The work opens the way to a new approach to the study of innovation, of what is possible and how this follows from what already exists.

The notion that innovation arises from the interplay between the actual and the possible was first formalized by the complexity theorist Stuart Kauffman. In 2002, Kauffman introduced the idea of the “adjacent possible” as a way of thinking about biological evolution.

The adjacent possible is all those things—ideas, words, songs, molecules, genomes, technologies and so on—that are one step away from what actually exists. It connects the actual realization of a particular phenomenon and the space of unexplored possibilities.

But this idea is hard to model for an important reason. The space of unexplored possibilities includes all kinds of things that are easily imagined and expected but it also includes things that are entirely unexpected and hard to imagine. And while the former is tricky to model, the latter has appeared close to impossible.

What's more, each innovation changes the landscape of future possibilities. So at every instant, the space of unexplored possibilities—the adjacent possible—is changing.

“Though the creative power of the adjacent possible is widely appreciated at an anecdotal level, its importance in the scientific literature is, in our opinion, underestimated,” say Loreto and co.

Nevertheless, even with all this complexity, innovation seems to follow predictable and easily measured patterns that have become known as “laws” because of their ubiquity. One of these is Heaps’ law, which states that the number of new things increases at a rate that is sublinear. In other words, it is governed by a power law of the form $V(n) = kn^\beta$ where β is between 0 and 1.

Words are often thought of as a kind of innovation, and language is constantly evolving as new words appear and old words die out.

This evolution follows Heaps’ law. Given a corpus of words of size n , the number of distinct words $V(n)$ is proportional to n raised to the β power. In collections of real words, β turns out to be between 0.4 and 0.6.

Another well-known statistical pattern in innovation is Zipf’s law, which describes how the frequency of an innovation is related to its popularity. For example, in a corpus of words, the most frequent word occurs about twice as often as the second most frequent word, three times as frequently as the third most frequent word, and so on. In English, the most frequent word is “the” which accounts for about 7 percent of all words, followed by “of” which accounts for about 3.5 percent of all words, followed by “and,” and so on.

This frequency distribution is Zipf’s law and it crops up in a wide range of circumstances, such as the way edits appear on Wikipedia, how we listen to new songs online, and so on.

These patterns are empirical laws—we know of them because we can measure them. But just why the patterns take this form is unclear. And while mathematicians can model innovation by simply plugging the

observed numbers into equations, they would much rather have a model which produces these numbers from first principles.

Enter Loreto and his pals (one of which is the Cornell University mathematician Steve Strogatz). These guys create a model that explains these patterns for the first time.

They begin with a well-known mathematical sand box called Polya's Urn. It starts with an urn filled with balls of different colors. A ball is withdrawn at random, inspected and placed back in the urn with a number of other balls of the same color, thereby increasing the likelihood that this color will be selected in future.

This is a model that mathematicians use to explore rich-get-richer effects and the emergence of power laws. So it is a good starting point for a model of innovation. However, it does not naturally produce the sublinear growth that Heaps' law predicts.

That's because the Polya urn model allows for all the expected consequences of innovation (of discovering a certain color) but does not account for all the unexpected consequences of how an innovation influences the adjacent possible.

So Loreto, Strogatz, and co have modified Polya's urn model to account for the possibility that discovering a new color in the urn can trigger entirely unexpected consequences. They call this model "Polya's urn with innovation triggering."

The exercise starts with an urn filled with colored balls. A ball is withdrawn at random, examined, and replaced in the urn.

If this color has been seen before, a number of other balls of the same color are also placed in the urn. But if the color is new—it has never been seen before in this exercise—then a number of balls of entirely new colors are added to the urn.

Loreto and co then calculate how the number of new colors picked from the urn, and their frequency distribution, changes over time. The result is that

the model reproduces Heaps' and Zipf's Laws as they appear in the real world—a mathematical first. “The model of Polya's urn with innovation triggering, presents for the first time a satisfactory first-principle based way of reproducing empirical observations,” say Loreto and co.

The team has also shown that its model predicts how innovations appear in the real world. The model accurately predicts how edit events occur on Wikipedia pages, the emergence of tags in social annotation systems, the sequence of words in texts, and how humans discover new songs in online music catalogues.

Interestingly, these systems involve two different forms of discovery. On the one hand, there are things that already exist but are new to the individual who finds them, such as online songs; and on the other are things that never existed before and are entirely new to the world, such as edits on Wikipedia.

Loreto and co call the former novelties—they are new to an individual—and the latter innovations—they are new to the world.

Curiously, the same model accounts for both phenomenon. It seems that the pattern behind the way we discover novelties—new songs, books, etc.—is the same as the pattern behind the way innovations emerge from the adjacent possible.

That raises some interesting questions, not least of which is why this should be. But it also opens an entirely new way to think about innovation and the triggering events that lead to new things. “These results provide a starting point for a deeper understanding of the adjacent possible and the different nature of triggering events that are likely to be important in the investigation of biological, linguistic, cultural, and technological evolution,” say Loreto and co.

We'll look forward to seeing how the study of innovation evolves into the adjacent possible as a result of this work.

Ref: arxiv.org/abs/1701.00994: Dynamics on Expanding Spaces: Modeling the Emergence of Novelties
