[INTERPLANETARY TRANSMISSIONS] [GENESIS]

Proceedings of
the Santa Fe Institute's
First InterPlanetary Festival

DAVID C. KRAKAUER
CAITLIN L. MCSHEA
editors

with illustrations by Caitlin L. McShea

Part I: Genesis

Part II: Proceedings of Santa Fe Institute's First InterPlanetary Festival

INTRODUCTION: SOCIAL & ECONOMIC ENGINEERING

Well, societies without a plan, that was history so far; but history so far had been a nightmare, a huge compendium of examples to be avoided. . . .

"And with our work," John continued, "we are carving out a new social order and the next step in the human story."

—KIM STANLEY ROBINSON, *RED MARS*

We humans are at an intersection. The convergence of ideas and events at this intersection is just coming into focus but is significant enough that the intersection might be, for human and possibly planetary history, a critical point—a "point" at which a small perturbation can cause the system to shift to a new state. A state that is, for example, either still organized, like ice—but in a different way than the current state—or disorganized, like a gas. This intersection might, more significantly even, be an origin point. Before we can stretch the limits of our collective imagination to consider what future history could originate from this moment in time, it is worthwhile to consider the nature of the convergence.

Machine learning and artificial intelligence have been around for a long time (that is, longer than the media hype would suggest)—since at least the 1950s, when, at the Dartmouth 1956 AI conference, the field is said to have been founded. The pace of research is now increasing exponentially, if the number of machine learning and AI papers uploaded to the physics arXiv is any indication.

The 1990s mark the start of the controversially named "big data" era on two fronts: the biological front, with high-throughput genomics data, and the "digital" front with the rise of personal computing (and, hence, the opportunity to track, record, and quantify individual behavior on a large scale) and individual interactions *vis-à-vis* social media. With these developments, we have available unprecedented microscopic data on interactions, behavior, genetics, and physiology, that can be harnessed to study the mapping between individual-level behavior and group or collective behavior, whether of cells, brains, or societies.

Coincident with these developments, the science of micro to macro in adaptive systems is blossoming due largely, I would say, to the rise of complexity science. The contributions of complexity science have been philosophical, conceptual, and technical. Philosophically, complexity science, borrowing from physics, presupposes that with the right lens it is possible to discover organizational principles that are scale- and substrate-independent. Conceptual contributions include moving away from substrate-specific questions to an emphasis on problems that reoccur across adaptive systems—for example, those relating to information processing and computation, robustness, communication and coordination, emergence, scaling, learning, and evolutionary dynamics. Technical contributions include techniques for studying micro to macro mappings by combining insights and approaches from statistical mechanics, theoretical computer science, and network theory; information theory for recasting evolutionary dynamics in terms of changes to mutual information; maximum entropy approaches and information theory for quantifying how collective or decomposable a system is (which is critical to a theory of control or intervention); and dimension-reduction and coarse-graining techniques to identify the dominant causal contributions to macroscopic change and hence build a theory for regularities observed at the macroscopic scale. One of the best examples we have so far of a successful identification and derivation of lawlike behavior in adaptive systems is the work on metabolic scaling of Geoffrey West and colleagues, which is accessibly discussed in West's book, *Scale*. As Freeman Dyson put it in a provocative but in some ways maddeningly incorrect review of Geoffrey's work and, more generally, SFI's research program, we (humans) are somewhere between Galileo and Newton in our understanding of adaptive systems.

Two other developments at this intersection are climate change, with its rapidly accelerating pace, and a re-injection of energy into space exploration—in particular, getting to Mars, largely through research and development by private companies like SpaceX, Blue Origin, and Virgin Galactic.

These five factors—AI, big data, a blossoming understanding of micro to macro in adaptive systems, climate change, and space travel—are linked, but not in a trivial way. A longer essay and a lot of thought would be required to work out the relationships. For now, the convergence is an observation with at least some clear if not yet concrete implications. With respect to this panel, the most relevant of these implications is that for the first time in human history a quantitative science of social, environmental, and economic engineering looks possible.

Humans have been attempting to engineer social outcomes since the dawn of cultural history. As I mentioned in the panel, there are many great examples. In his book *Priests and Programmers*, anthropologist Steve Lansing describes how a Balinese water temple system emerged in the ninth century to optimize planting cycles and water distribution. A Rube Goldberg–like voting process for electing the Doge, described beautifully by John Julius Norwich in *A History of Venice*, was invented somewhat cooperatively by rival Venetian families in the 1500s to help prevent the process from being gamed.

These examples, however surprisingly elegant, differ from canonical examples of engineering and manufacturing plants and cars and spaceships in that there are no blueprints for social systems, no rigorously quantitative way yet to identify targets of interventions that will reliably produce or control change. Rather, the history of human social, financial, and ecosystem engineering is based largely on intuition. As Isaac Asimov wrote in an essay in *The Planet that Wasn't*,

> "People are entirely too disbelieving of coincidence. They are far too ready to dismiss it and to build arcane structures of extremely rickety substance in order to avoid it. I, on the other hand, see coincidence everywhere as an inevitable consequence of the laws of probability, according to which having no unusual coincidence is far more unusual than any coincidence could possibly be."

Perhaps consequently, the majority of attempts to engineer adaptive systems have been disastrous or impotent, especially those that did not have the benefit of developing organically over a long time period, as self-organization and long timescales can sometimes compensate for cruddy intuition. We might call the past history of social engineering reactive. The future can in principle be proactive.

Humans have been attempting to engineer social outcomes since the dawn of cultural history.

The behavioral maps we will be able to build with the vast microscopic data now being collected, developments in AI and complexity science, and motivation from the desire to get to Mars and control the climate might allow us to find and quantify the hidden regularities in our social interactions—regularities that, thus far, we have been unable to measure, or which may have been invisible given our myopic perception and intuition-dominated reasoning. Not to mention the fact that most of the adaptive systems we want to influence are complex, with multiple time and space scale, heterogeneous actors, and learning, as well as evolutionary dynamics.

If we can successfully infer the rules and strategies individuals use to guide decision-making, we will have a robust starting point for building predictive simulations of social outcomes at the societal level. Such simulations will also allow us to test alternative futures and give a quantitative, empirical basis for our intervention decisions.

The potential power of this approach seems obvious. One only has to look at the huge investment into data collection by corporations like Google and Facebook and all of the third-party companies—data merchants—that deal solely in data sales, or to China's social credit program. But, amazingly, just ten years ago, there was little discussion outside of science-fiction novels of this growing reality. The public as well as many scientists scoffed at it as a pipe dream. The 2016 Facebook election debacle—even if it is overhyped—is an example of just how poorly we collectively anticipated the change from reactive to the beginnings of a proactive, quantitative social engineering, and how rapid its initial stages might be.

Will there be a giant leap forward in proactive social engineering allowing the orchestration of precisely engineered individual- or societal-level outcomes? Probably not. Two reasons why this is unlikely are the stochastic (random), rather than deterministic, nature of human behavior and the stochasticity in the process by which behavior combines to produce society. Even if scientists had the best data and methods at their disposal, complete prediction would never be possible because of the character of adaptive systems. Adaptive systems are error-prone computers making estimates based on finite, imperfect data, and they are subject to changing environments. One might respond to this pessimism by suggesting that social engineering could reduce behavioral variance and hence eliminate much noise. But that view is naïve, and brings us to a second reason why social engineering will never rival standard engineering in its predictive power.

Adaptive systems are just that—adaptive; their actors respond strategically as the system and environment changes, in evolutionary or learning time, meaning that metrics used to gauge adaptation, once they become targets, can cease to be useful metrics. This is Goodhart's Law, and it applies when the timescale separation between the microscopic and macroscopic is too small. Social engineering may in fact make social systems more, not less, complicated and predictable if these engineering arms races get out of control. Perhaps a way around this is for social engineering to focus on process over orchestrating specific outcomes like the degree of inequality.

What is much more likely than tight control over the future is coarse but robust prediction based on an understanding of dominant causes at the mesoscale. This is illustrated by a recent study in *Science* by Nicolas Bain and Denis Bartolo of the collective motion of marathoners. This study found the large-scale motion of the runners could be predicted without knowledge of individual interaction rules if the crowd was modeled as a fluid. The question for social engineering is: can we build societies that are like fluids, so that we can predict and control aggregate behavior without having to know, or care about, what the individuals are doing? Assuming that such models work even when the range of individual behavior is large and varied (unlike in marathons), this would resolve many problems concerning individual–collective trade-offs, such as privacy and autonomy, and might allow a societal engineering to which we can all contribute cooperatively and adversarially, as our local needs dictate.

A little imagination has gone into this introduction. If we are to build an interplanetary civilization and want to make use of our growing capacity for social engineering to do it, we are going to need a lot of thinking outside the box. The purpose of the InterPlanetary Festival and this panel in particular is to accelerate that discussion.

—Jessica Flack
Professor, C4 Director, & Chair of Public Events
Santa Fe Institute