Profile of Michael I. Jordan

The multidisciplinary work of Michael I. Jordan, a recently elected member of the National Academy of Sciences and the National Academy of Engineering, is perhaps best introduced by what he calls one of the most fascinating intellectual questions of all time: How does a person reason and make decisions in environments that are uncertain? "Humans are in a world where most of the entities that they care about are not ever explained to them," says Jordan, professor of Statistics and of Electrical Engineering and Computer Science at the University of California at Berkeley. "The world is composed of a sensory stream and somehow we have to figure out what parts of that sensory stream are important, what parts have semantic relevance, and what parts relate to our goals and our needs. Even if we learn all about language and we learn about the meanings of words and are able to perceive and manipulate objects in the world, we still are uncertain," Jordan says.

That uncertainty arises in part because humans have only a partial knowledge about the world and are ignorant of many influential factors operating behind the scenes, but also because the world is inherently stochastic. "How does an intelligent entity cope with the vast amount of uncertainty around him or her?" Jordan wonders. "And how can we understand that, how can we make it better, to improve our lot in life?"

To explore these questions, Jordan has embarked on an intellectual odyssey that has taken him across the far reaches several fields, including psychology, statistics, cognitive science, computer science, and engineering. His work has helped to forge new links between these fields, and has provided a deeper understanding of the relationships among learning, inference, induction, and reasoning.

Exploring the Unknown

A late baby boomer child of the 1960s, Jordan's childhood was divided between Louisiana and Kansas as his parents moved back and forth. He cultivated a curiosity for exploring the unknown by reading the tales of Renaissance explorers, such as Marco Polo, whose epic journeys across Asia led him to discover new places and to learn about new cultures. Having witnessed the social movements of the 1960s and 70s, Jordan was inspired to do something that had to do with human beings and social phenomena, and his intellectual interests soon gravitated toward exploring the largely unknown inner workings of the human mind, intelligence, and psychology.

Michael I. Jordan.

Captivated by the vibrant culture, food, and music of Louisiana, Jordan decided to attend Louisiana State University, where he majored in psychology. Jordan remembers the first time he entered the campus's large library, being shocked by the huge number of books on all kinds of things about which he knew nothing. "I decided at that moment that I'd like to learn a significant fraction of what was in the library," he says. Jordan began reading the work of the mathematically oriented philosophers Bertrand Russell and Kurt Gödel, and developed a keen interest in understanding how humans think. "It was clear to me that that psychology was going to become a more scientific field, and I wanted to be a part of that," Jordan recalls.

Adventures in Math

After completing his bachelor's degree in 1978, Jordan decided to work toward becoming a mathematical psychologist and began a master's program in mathematics and statistics at Arizona State University. "Since I was going into psychology, which is an experimental field, you needed to learn how to analyze the data, so you needed to take the classes in statistics," Jordan explains. However, Jordan soon realized that he didn't want to learn statistics just so he could analyze data, but so that he could build new models and explore how statistical inference is related to human thinking. While on vacation from Arizona in San Diego, Jordan visited the campus of the University of California at San Diego, and discovered the emerging field of cognitive science after meeting with some of the faculty there. "I had a kind of 'Aha!' experience that this was for me," he says. "It was, in that era, a new field, and it was really about math and science applied to how humans think. That inspired me."

Jordan moved to San Diego after completing his master's degree in 1980, and shortly thereafter, began working on his PhD in cognitive science with David E. Rumelhart, a member of the National Academy of Sciences and a professor of psychology and cognitive science at the University of California at San Diego. "He was a major figure in the field, and he was a mentor for me," Jordan says. "He helped me to deepen my interests in learning and reasoning, and mathematical approaches to how people think." Jordan was initially drawn to Rumelhart's work on natural language semantics: the study of the meaning of words and sentences, and how humans learn to understand what words and sentences mean. However, by the time Jordan arrived there, Rumelhart was focusing on two other compelling topics. One topic was the nascent area of neural networks, and involved making mathematical models of brain-like computation to better understand human reasoning, decision-making, memory, and language use. The other area of research involved making models of human movement to better understand the underlying neural control system that humans use to move their limbs. Jordan's initial focus was on the latter topic, and he set off on a program of research involving both experimental work and modeling. His modeling work aimed to exploit the rich tradition of control theory, the mathematical field that aims to analyze and design self-regulated systems. The central nervous system of the body can be considered as a type of controller that receives sensory inputs from the environment and dictates the appropriate motor response.

"I was interested in control theory because I was trying to understand how humans make movements, how they would then adapt their movements if they're making incorrect movements, and in general how humans learn to move so as to manipulate objects in the world," Jordan says. Control theory is based on building models of the dynamic system that is being controlled, using the kinds of

This is a Profile of a recently elected member of the National Academy of Sciences to accompany the member's Inaugural Article on page 1160.

statistical modeling ideas with which Jordan was already familiar. He then used optimization theory to extract from the model the appropriate control signals that are needed to achieve desired goals. "This blending of statistics and optimization continued to characterize my work long after I had moved on to other topics outside of motor control," Jordan says.

Turning to Machines and Starting a Laboratory

Toward the end of his PhD training, Jordan had begun to feel a need to return to broader issues in inference and decision-making, and sensed that his statistical and control-theoretic perspectives needed to be augmented by the algorithmic perspective provided by the field of computer science. Thus, after completing his PhD in cognitive science in 1985, Jordan went to the University of Massachusetts at Amherst to work as a postdoctoral fellow in the field of artificial intelligence with Andrew Barto. "When I started my work in artificial intelligence, it seemed like there was a lot of work on reasoning, but not much work on learning," Jordan recalls. "One of the thoughts I had was that reasoning and learning can aid each other, so you don't have to learn what you can reason about, and learning can avoid having to do complicated reasoning. These are two sides of the same coin of reducing uncertainty, but they're pretty different in terms of the way you interact with data and the kind of algorithms you talk about. So as time went on, I tried to focus more and more on algorithms for learning and not just how humans learn, but how machines can learn as well."

After two years of postdoctoral research, Jordan accepted an assistant professor position in the Department of Brain and Cognitive Sciences at the Massachusetts Institute of Technology in 1988, and began setting up a research group that focused both on human motor control and on machine learning and reasoning. In human motor control, Jordan's team (consisting of graduate student Zoubin Ghahramani and postdoctoral fellow Daniel Wolpert) began to study how people make reaching movements as they adapt to a distorted visual scene. "We used a device such that when humans would look at the device, they would not see the truth; they would see a perturbed version of the truth," Jordan says. Under the influence of this altered visual scene, the study participants would reach for an object and fail to do so. However, as time passed, the study participants would learn how to behave in the altered visual environment and correct their movements. Jordan's team measured the movements, and used this experimental data to make models of what might be happening in the brain as people learn to make movements (1–3).

Jordan also performed an analogous set of experiments with graduate student John Houde to study people's ability to adapt to altered speech production (4). "We built a little device that would plug into people's ears so that they couldn't hear themselves speaking, and then we would take the sounds that would come out of the microphone, and we would change the sound so that when a person said 'ah' it would sound more like 'oh'," Jordan explains. The participants, wearing headphones that played back only the altered versions of the particular sounds they made, learned to adjust the shapes of their mouths so that the sounds would come out a bit differently, until ultimately succeeding in creating the sound that they had intended to make. Studying this adaptive process under various probes allowed Houde and Jordan to test hypotheses about the control units underlying human speech production.

Exploring Statistical Research

On the mathematical and algorithmic front, Jordan made use of his background in optimization theory to begin exploring a variational perspective on statistical inference (5). Within this perspective, the solution to a statistical inference problem is expressed as the solution of an optimization problem, and this optimization problem is perturbed, or "relaxed," so that it can be solved more readily. Jordan was also engaged in research in the field of probabilistic graphic models, where Judea Pearl and others were making use of graph theory to solve problems in reasoning, and Jordan perceived that the variational approach to inference fit naturally with the graphic model framework. The end result was an algorithmic toolbox in which graph-theoretic algorithms could be used to combine reasoning and learning in the ways that Jordan had envisaged. This overall framework has seen a variety of applications in scientific and engineering problems, including the study of computational biology, speech, natural language, vision, robotics, and information retrieval.

In 1997, Jordan was contacted by Peter Bickel, a statistician at the University of California at Berkeley, who invited Jordan to join the University of California's statistics department as well as the computer science department. "That sounded like a wonderful learning experience and an intellectual challenge," Jordan says, "so I took the opportunity and moved to Berkeley."

At the University of California at Berkeley, Jordan dove head first into mathematical and computational issues. Continuing in the graphic model and variational inference vein, he and his students developed a model known as Latent Dirichlet Allocation (LDA), which aims to discover the topics underlying collections of documents (6). LDA is an example of a so-called latent variable model, where one envisages a simple underlying description for a complex observed object and builds a model and inference procedure that allows one to recover that underlying description from data, Jordan says. "For example, I might give you a document out of the *New York Times*, and the document is about sports, but the word 'sports' never occurs. So it's not a keyword associated with the document. But you'd notice that there are lots of words that would be likely to be in this document if sports were the underlying topic that the writer had in mind." LDA enabled the discovery of this latent, or hidden, variable: in this case, that the document is about sports. Moreover, LDA allowed documents to exhibit blends of topics, a key capability that allowed LDA to find a range of applications beyond the original application to documents. "This 'wide range of applicability' is what I find particularly appealing about research in statistics," says Jordan. "And I'm also pleased to note that there have been many applications of LDA to problems in cognitive science."

As Jordan dug further into statistics and computer science, he began to miss pure science. He soon got involved in a number of collaborations in which he was able to bring his methods from statistics and computer science to bear on modern molecular biology problems: phylogenetics, genomics, and proteomics. "These areas in molecular biology are generating lots of data," Jordan says, "and they have lots of underlying latent variables they'd like to make inferences about." Jordan, who thinks of his statistical methods as a kind of toolbox, has since used his "tools" in creative ways for each unique problem he encounters. He has helped to uncover the ancestral populations underlying an organism's genome, to analyze proteomic, transcriptomic, and genomic profiles, and to predict various functional and structural properties of biological molecules, such as proteins and DNA (7–15).

Jordan's Inaugural Article (16), coauthored with his former graduate student Alexandre Bouchard-Côté, outlines a method to simultaneously align a set of related genetic or protein sequences and infer the sequences' phylogeny, a challenging problem in comparative genomics. "Nowadays, the data in phylogenetic analysis are often DNA or other strings, and there have been deletions and insertions over evolutionary history," Jordan says. "If you take two organisms and look at their DNA strings, you don't know a priori what DNA aligns with what other DNA." The solution, Jordan says, is to align those strings and build the phylogenetic tree at the same time. "Lots of classic work in phylogenetics assumes the alignment is already given and then builds the tree. And most algorithms that do alignment assume a tree has been given, and then given the tree, builds the alignment."

However, this type of approach is biased, Jordan says. "If two organisms have the same insertions and deletions over evolutionary time, then they're likely to be closely associated in the phylogenetic tree. So if you infer the insertions and deletions as part of the tree building, you're able to take that into account. If on the other hand you assume the alignment is finished, then the tree can't take into account the insertions and deletions, and it's not as good of a tree; it's statistically biased."

Relentless Curiosity

In his free time, Jordan attempts to satisfy his intellectual curiosity by reading voraciously and learning new languages. A native speaker of English, he has also learned to speak Spanish, French, and Italian, inspired by the marvelous ability of the human mind to learn new things. "Even though no one is really teaching me—I'm just sitting down and reading a novel in another language and looking up words in another dictionary—but somehow I'm learning all these subtle things in another language," he muses. "How is that possible?" In the course of learning the languages, Jordan also professes to have learned a lot about the cultures behind the languages, which includes their music and their food. "I'm a very curious person, so I like being able to open doors and to access people and traditions and cultures that I wouldn't have otherwise known about."

He also plays many different musical instruments and is currently focusing on the drums. "I love the physical motor-control side of playing the drums, training the body to make music with the limbs," he says. When a person plays the drums, all four limbs are often involved in making sounds. Because the feet are farther from the brain than the hands, for these limbs to strike the drums simultaneously means that the brain sends signals to the feet earlier than the signals to the hands. "We don't have to think about this, but our brains somehow just figure it out magically. It's fascinating to me that that's all being worked out. I think these are open scientific problems for the next several generations of trying to understand how the brain is aware of what's around it, and what controls its limbs and controls entities in the world to make desirable consequences happen."

Nicholette Zeliadt, *Science Writer*

1. Wolpert DM, Ghahramani Z, Jordan MI (1994) Perceptual distortion contributes to the curvature of human reaching movements. *Exp Brain Res* 98(1):153–156.
2. Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. *Science* 269 (5232):1880–1882.
3. Ghahramani Z, Wolpert DM, Jordan MI (1996) Generalization to local remappings of the visuomotor coordinate transformation. *J Neurosci* 16(21):7085–7096.
4. Houde JF, Jordan MI (1998) Sensorimotor adaptation in speech production. *Science* 279(5354):1213–1216.
5. Jordan MI, Ghahramani Z, Jaakkola TS, Saul LK (1999) An introduction to variational methods for graphical models. *Mach Learn* 37(2):183–233.
6. Blei D, Ng A, Jordan MI (2003) Latent Dirichlet allocation. *J Mach Learn Res* 3:993–1002.
7. Bouchard-Côté A, Sankararaman S, Jordan MI (2012) Phylogenetic inference via sequential Monte Carlo. *Syst Biol* 61(4):579–593.
8. Sankararaman S, Kimmel G, Halperin E, Jordan MI (2008) On the inference of ancestries in admixed populations. *Genome Res* 18(4):668–675.
9. Engelhardt BE, Jordan MI, Srouji JR, Brenner SE (2011) Genome-scale phylogenetic function annotation of large and diverse protein families. *Genome Res* 21(11): 1969–1980.
10. Blum B, Jordan MI, Baker D (2010) Feature space resampling for protein conformational search. *Proteins* 78(6):1583–1593.
11. Wauthier F, Jordan MI, Jojic N (2011) Nonparametric combinatorial sequence models. *J Comput Biol* 18(11): 1649–1660.
12. Ting D, et al. (2010) Neighbor-dependent Ramachandran probability distributions of amino acids developed from a hierarchical Dirichlet process model. *PLOS Comput Biol* 6(4):e1000763.
13. Flaherty P, et al. (2008) A dual receptor crosstalk model of G-protein-coupled signal transduction. *PLOS Comput Biol* 4(9):e1000185.
14. Gyaneshwar P, et al. (2005) Lessons from *Escherichia coli* genes similarly regulated in response to nitrogen and sulfur limitation. *Proc Natl Acad Sci USA* 102(9):3453–3458.
15. Giaever G, et al. (2004) Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc Natl Acad Sci USA* 101(3):793–798.
16. Jordan MI, Bouchard-Côté A. (2013) Evolutionary inference via the Poisson Indel Process. *Proc Natl Acad Sci USA* 110:1160–1166.