



## *Playing “20 Questions” with Geometry*

# ROCCO A. SERVEDIO

Associate Professor of  
Computer Science

In “20 Questions,” one player thinks of an object and the others get 20 yes-no questions to guess its identity. “That’s easy, but what if you let the answerer lie three times? That makes it much more difficult,” Rocco Servedio said.

That is the type of problem researchers face when false signals, or noise, corrupt data. Servedio’s goal is to develop robust algorithms that learn complicated rules even in the presence of noisy data. Such algorithms could learn patterns that improve sensor performance, predict earthquakes, or forecast financial markets.

One of Servedio’s most powerful tools is geometry. “When you cast a learning problem in a geometric framework, you’re often on the way towards solving it,” he said.

Imagine, for example, a piece of paper with red plus signs and green minus signs on opposite sides of an unknown dividing line. A few pluses are mixed with the minuses and vice versa. “In this two-dimensional example, you can eyeball the data and see which points don’t belong. In higher dimensions, where each point has many coordinates, this is much more difficult, though we can sometimes pull it off with tools from high-dimensional geometry,” he said.

“The way people understand something is by drawing pictures. I’m usually working in high-dimensional Euclidean spaces where it’s tough to draw accurate pictures,” he said, “but thinking geometrically still provides useful insights.”

Servedio also takes a geometric approach to studying rules used to classify information. One popular approach is the decision tree. Like “20 Questions,” it uses a sequence of yes-no questions to decide how to label data points.

“If you think of this logical representation geometrically, you can sometimes see properties that would have otherwise remained hidden. These insights can lead to better learning algorithms,” Servedio said.

Servedio also uses geometry to compensate for missing data. Imagine that it takes 1,000 coordinates to describe a data point completely. What kind of learning is possible if only one of those coordinates is available?

“There are ways to compensate for massive amounts of missing data,” Servedio said. “It might sound impossible, but doctors do something like this all the time. They could potentially run thousands of clinical tests on a patient to fully describe his or her condition, but a good doctor can make a useful diagnosis from just one or two tests.”

*A.B., Harvard, 1993; M.S., Harvard, 1997; Ph.D., Harvard, 2001*