One Hundred Thousand Questions

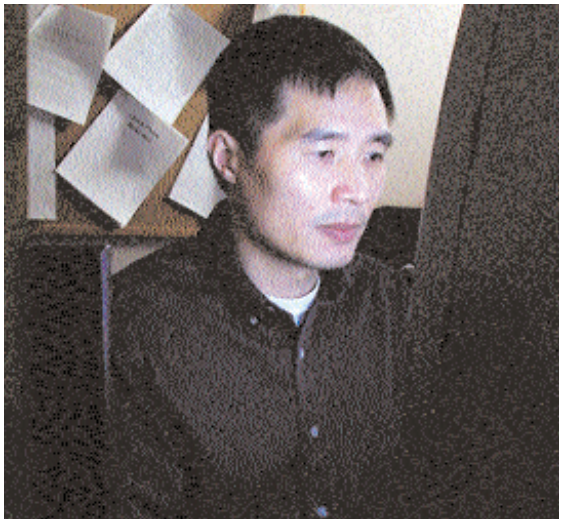
In 1970, C. P. Yang's father, Xitian Yang, returned to his wife and two young sons in Guangzhou, China, after two years in a labor camp, a toll extracted from many intellectuals during China's Cultural Revolution. His wife, mindful of the potential danger that her upper-class background presented to herself and her family, continued her work in a factory. However, the Research Institute of Building Materials, which Xitian had once headed, was closed and he was not permitted to work. That gave him ample time to devote to both his systematic mastery of five languages and the proper education of his sons. Little could Xitian imagine that his efforts would bear fruit in one son's career as a petroleum engineer in Houston and C. P.'s success as a U.S. physicist who would craft the scientific graphing and analysis software used by more than 80,000 scientists and other researchers around the world.

Other than providing ample spare time, China at the time could not have been a more inhospitable environment for Xitian to pursue his goals because nearly all scientific journals, texts, and books were banned. But somehow Xitian found loopholes that brought stacks of contraband scientific periodicals from abroad into the Yang home. C. P. and his brother, A. P., watched and learned how to turn mountains of index cards into self-study tutorials that could unlock the languages and the treasures of these scientific works.

Education

The family also had a few banned science books, including a couple of volumes of *100,000 Questions*, a question-and-answer encyclopedia of science that had been a highly influential staple of science education in China before the Cultural Revolution. Xitian swapped their few volumes for others so that the two Yang boys could read nearly all the questions.

"When I was in elementary school, my father showed me the 100,000 questions book," he recalls. "It was not just about physics but all science subjects, and it probably had a great deal to do with why I wanted to be a scientist."



During China's Cultural Revolution, C. P. Yang's father managed to obtain some volumes of the banned science book *100,000 Questions*, which C. P. believes was influential in his later decision to become a scientist.

But planning for a career in the sciences was not an option for the young C. P. The Cultural Revolution began in 1967, the year he entered elementary school, and continued for the duration of his high school education. As a student during that period, Yang was expected to go to work in a rural area after his graduation. Luckily, Chinese leader Deng Xiaoping changed that policy in the mid-1970s, and Yang attended the University of Science and Technology in Hefei, where he majored in geophysics. He was not the best student in the class, perhaps in part because of his continued devotion to reading the various books and journals that came his way in lieu of studying assigned texts.

Reading about physics and other sciences, sometimes in English, paid off, however, when C. P. scored high on the Chinese and United States Physics Examination and Application test, which led to a scholarship for physics graduate work at Carnegie Mellon University beginning in 1982. There, Yang attained his master's and doctoral degrees in biophysics and devoted himself to studies of model membranes. During this period, he had his first encounter with what later became his ardent passion, computer programming.

"When I applied to Carnegie Mellon, I actually had no idea that the school had a well-deserved reputation for excellence in computer science," Yang says. "My experimental work constructing membranes using lipid bilayers involved a lot of programming because at that time there weren't many tools available for data analysis or graphing. Instead, I used tools that I developed in Fortran for DOS."

The software provided basic data-analysis functionality, including data smoothing, integration, and statistical analysis. It became popular enough at Carnegie Mellon that the university installed it at the computing center and provided for its support through the academic computing department.

"My training as a physicist had bearing on the solutions I developed then, as it does now," Yang says. "I realized that what was needed was a way to draw a graph that could be easily manipulated to overcome the difficulties of managing a graph with multiple axes. Most software then did not address the problem this way, and multiple axes led to complications. Instead, I looked at graphs as a physicist looks at graphs, that is to say, as a coordinate system where the ability to control the axes is key."

Origin

Upon completing his doctoral studies at Carnegie Mellon in 1988, Yang went to City College of New York for postdoctoral work, where he developed another data-analysis system in C for Macintosh. Word of his programming achievements reached John Brandts, a chemist by training and founder of MicroCal, Inc., which manufactures calorimeters. Brandts was seeking someone

Creating the Next Generation

When C. P. Yang and his co-workers at OriginLab sit down to plan the next generation of Origin software, the brainstorming often begins by revisiting the enduring challenges that physicists face when interpreting experimental data. First, that means ensuring that physicists and other scientists can not only graph their data, but also have complete control over all aspects of their graph and analysis procedures. This creates a general imperative for ease-of-use throughout the software.

The trick is to find elegant programming solutions that will be advantageous to researchers using diverse experimental methods and techniques. Years ago, for example, Yang was the first to apply the concept of double-click editing to PC-based scientific software to enable a user to modify a graph simply by double-clicking on the element to be changed and then graphically selecting new formatting options.

The earliest Origin software also originated a way for physicists and other users to plot data on multiple layers. This was the first time one could define layers of axes within graphs, and it enabled an intuitive way to construct the most effective viewing of data. Today's programming innovations include building the most common programming language, C/C++, into the Origin graphing software so that users do not have to learn a new language to automate their analyses. The built-in C programming lets scientists create their own mathematical algorithms, as their research requires.

"Researchers each have their own unique requirements for data-analysis software, yet they have no spare time for learning to use new software," Yang says. "Therefore, we have focused on making Origin so easy to use that you will be able to produce results right out of the box. In addition, we have put a C compiler into the software so that one can perform serious scientific computation using a familiar language."

to write a data-analysis program for users of MicroCal instruments. Although Microsoft Windows was a fledgling and little-known operating system in 1989, Yang envisioned that it would allow the development of the most user-friendly application. His solution became a strong asset for MicroCal, and the company decided in 1991 to release the software to the public as a separate product. Its release roughly coincided with Windows' rapid penetration of the market. The software, named Origin, was only one of two Windows-based data-analysis applications commercially available at that time. Seeing a business opportunity, Brandts and Yang formed a spin-off company, Microcal Software, to more aggressively develop and market Origin. In 2000, Microcal Software was renamed OriginLab.

On a personal level, creating Microcal Software catapulted Yang from academia into the challenges of entrepreneurship.

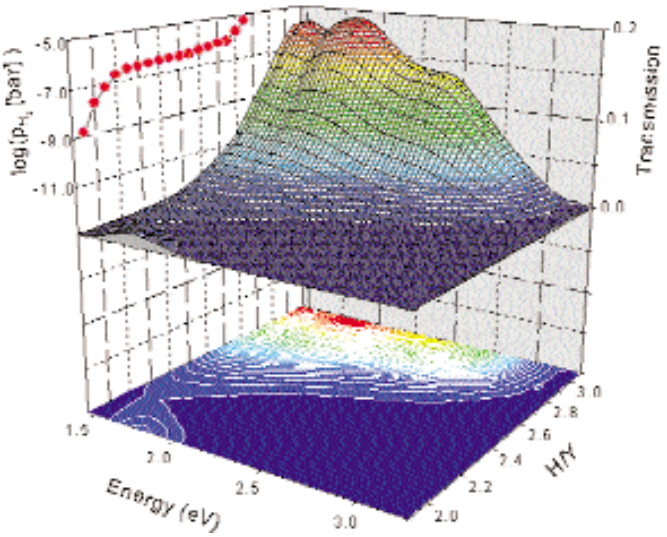
"It was very difficult in the beginning," C. P. recounts. "In the academic world, I enjoyed funding within a certain framework and mainly had to manage within those boundaries. But there is no such certitude for an entrepreneur, and I had to adjust to the fact that even though I would like to create something fun and interesting, I must also be mindful of the bottom line. On the other hand, I believe that coming from academia also gave me the advantage of having a longer-term vision. It takes a tremendous amount of determination to carry through for the company in the long run. I was lucky because John Brandts had already started a company in the early 70s and knew how to help us recover from the typical missteps of a start-up company."

Guangzhou

During the course of an average workday, Yang spends much of his time involved in instant messaging "conversations" with his staff in Boston (where Origin's core code is developed); Northampton, Massachusetts (the company's headquarters); and Guangzhou, China (where each day's newly developed code is tested overnight).

Managing a company on opposite sides of the globe presents both challenges and possibilities. Yang calls the China-based testing team an indispensable part of the Origin development process. "We set up the Guangzhou office in the middle of three universities so that OriginLab could provide an opportunity for students to work in an entirely English environment devoted to software testing," he says. "Most of the students are doing graduate work in physics, mathematics, and related fields. We use the Web to access our bug-tracking database so that all of the testers in China and all of the code developers in the United States communicate in a group environment through the tracking database. All these communications are in English, and while all may not be entirely in perfect English, we make sure everyone communicates clearly."

For the students, the opportunity to both improve their English and work for a U.S. software company is an attractive lure that works well for OriginLab. In fact, more students want to work for the company than it



The transmission spectrum of a 30-nm yttrium hydride film can be plotted in a 3-D color map as a function of hydrogen concentration in the film (top), along with a contour plot of energy levels (bottom) and a pressure–composition isotherm (left).

can hire in its Guangzhou office.

Because Origin is graphical-user-interface oriented, the software requires labor-intensive testing. Every area of the software is tested repeatedly in different language environments (including English, German, and Japanese) and in different operating systems (Win98, WinME, WinNT, Win2000, and WinXP). Because code testing is highly repetitive, a large number of students perform the task on a rotating schedule. Unlike their counterparts in the United States, Chinese graduate students do not carry the heavy burden of teaching duties. Thus, they have more time to work for companies such as OriginLab that are near universities and offer flexible schedules.

"There is a huge commercial advantage to rapidly testing our code, which the Guangzhou office facilitates," Yang says. "Origin is a complex software, and any initial change made in the code will often lead to problems in completely unexpected places. Like other software developers, we expect that any time we fix a bug, there is a 30% chance that we will introduce another bug. But with the help of the team in Guangzhou, we are always testing the latest code so that our programmers coming into work in the morning will know that the bug report on their screen is about the code they put in the night before."

Physicists

Other than testing, all OriginLab functions are handled by the two U.S. offices, including development, support services, sales, and marketing. Like Yang, eight of the U.S. staff members are physicists, several of whom also emigrated from other countries—Russia, England, India, and the former Czechoslovakia. Collectively, their prior physics research—in theoretical particle physics, magnetism in intermetallic compounds, superconductors, spectroscopy, astrophysics, and medical image processing—brings a broad scope of expertise to bear on Origin's development and corporate culture.

The physicist-heavy staffing of Origin Lab is not accidental. "Although we develop tools for scientists in a wide range of disciplines, we always seek the common denominator in terms of how these disciplines deal with math and underlying functions," Yang says. "We find that training in mathematics or physics is required to construct such complex solutions."

Beyond his daily duties of conferring with the Origin development teams and other staff, Yang maintains evenings and weekends as his preserve for programming. "I feel a natural bond with programming and realize that computers can do much more so that scientists can focus on what they need to focus on," he says. "At this point, we are just scratching the surface of what the computer is capable of doing for scientific research." Ω

BIOGRAPHY

Joe Przechocki is product marketing manager for OriginLab Corp. in Northampton, Massachusetts (joe@originlab.com).