



Cool Programmer Challenge: Football Algorithm = \$50,000

By Esther Schindler

September 16, 2008

Imagine that you have detailed data from every college football game from the last four years, and you want to predict which teams will win through the rest of this season. How would you write a program to get the most accurate answer?

Writing software to identify the winning teams isn't a matter of a computer language choice or development methodology; it starts with the algorithm. Programmers can't start writing code until they identify the application's logic. An algorithm encapsulates that logic and essentially is a step by step recipe for solving the business problem. If the algorithm is faulty, the software is inefficient, inaccurate or otherwise fails in its goals. You probably learned that in Computer Science 101.

Most discussions about optimizing and innovating new algorithms, however, are—well, frankly, they are pretty dull. While it might be important to do [data mining](#) and to [perform data analysis](#) on plain old corporate data, occasionally something comes along that's a whole lot more fun. Certainly, the challenge offered right now by [ESPN](#), run by [TopCoder](#), should attract the attention of plenty of software developers—because of the prize money, if nothing else. (This could, however, bring closure to any lingering geeks-versus-jocks high school issues for all involved.)

It is a pretty problem, both on business grounds and as a technical puzzle. "We are trying to write an algorithm to predict the future outcome of college football games based on past performance," explains TopCoder's Bill Atwood, the TopCoder project manager for this ESPN project. ESPN plans to use the algorithm for prognostication, on-air prediction and pregame previews. That's a competitive advantage for ESPN, which can use accurate predictions to drive more viewers to their TV channels and website, points out Atwood.

In other words: as fun as this project might be, it has real business implications and could as easily be applied to duller IT tasks (though probably without as much programmer enthusiasm).

As is explained in depth on [the ESPN site](#), developers are given a huge amount of data to work with: four seasons of every college football game on a play by play level.



There are four phases to the competition. A preview has been underway for the first part of the college football season, during which TopCoder developers have already been hard at work. "Over the first two weeks we had fantastic results," says Mike Morris, TopCoder's senior vice president of software development. The top 10 people in the pool picked 77 percent of the winners, which is right at the top of the Las Vegas odds, according to TopCoder. "In the first week, we predicted the [UCLA-Tennessee](#) upset, where Tennessee was the 7.5 point favorite," brags Morris. "In the second week, Maryland was predicted to beat Middle Tennessee by 13 points, and we had that spot-on."

To read how the [National Football League](#) uses optimization software and mathematical models to make its complex schedule, see "[NFL Schedule, Rivalries and Potential TV Ratings Optimized by Packaged Software.](#)"

Now the developers in the contest are tweaking their algorithms for the official scoring.

The intriguing business model here is what Atwood calls "the wisdom of the best." People who can invent a brilliant algorithm may or may not be the ideal day-to-day programmers; and you might need a brilliant algorithm only once a year. After the development project is underway and it's all just code slogging, the algorithm goddess might no longer be necessary. A contest like this lets the enterprise pay for only that single big answer. (In other words: "Here's the best design—now I'll be on my way.")

The algorithm contests that TopCoder runs—this is by no means the only one—tap into this very specific niche, says Atwood. And since developers compete to write the best algorithm ("best" measured with metrics like "most accurate data prediction" in this instance) the client can have confidence that the algorithm they pay for is a good one. It is, as Atwood explains it, a really efficient model.

It's also a compelling one for developers. The first-place prize is \$50,000, for the best algorithm, which is based on everything after the initial two weeks of the college football season. The total prize package is \$100,000, with staggered prize money for those who come in second through fifth. (*I would stagger if I won.*)

It is not, however, a task that only a football fan can get into. The top three contestants on the leader board are from Poland, Egypt and Ann Arbor, Michigan. To help developers learn how American football works, such as how games are scored, ESPN's Will Harris is writing a Football 101 blog with game summaries, sharing advice and guidance on the sport. TopCoder guys are handling technical summaries about the data organization.

© 2008 CXO Media Inc.