

Control

continued from page 8

Intelligorithms

Dimitri Bertsekas and John Tsitsiklis, electrical engineers in the Laboratory for Information and Decision Systems at MIT, study applications of dynamic programming to large-scale stochastic control problems, such as those occurring in the scheduling and queueing theory. Together with Benjamin Van Roy, also of MIT, they have developed algorithms that combine complex feature extraction and simple approximations to cope with the otherwise overwhelming dimensionality of the "state space" that such problems entail. (Just to specify whether each of 100 queues is occupied or not, for example, gives rise to an unmanageably large number of states.) Their algorithms are then adjusted during a "training" period. One of the important theorems Tsitsiklis and Van Roy have proved is that these adjustments are guaranteed to converge to a "good" algorithm—something that various other approaches only purport to do.

Their results, Tsitsiklis says, build on methods pioneered by Andrew Barto and Richard Sutton at the University of Massachusetts at Amherst, which have led to the development of a world-class backgammon player by Gerald Tesauro at the IBM T.J. Watson Research Center. The MIT researchers "saw great promise in these methods," Tsitsiklis says, but also felt the subject needed "a solid mathematical foundation to provide guidance and insight into the merits of different methods." Their theoretical work has combined tools from dynamic programming, function approximation, and stochastic approximation. (Neuro-Dynamic Programming, Bertsekas and Tsitsiklis's recent book on the subject, opens with a quote from Confucius: "Learning without thought is labour lost; thought without learning is perilous.")

At the heart of any dynamic programming approach to stochastic control is what's called the "cost-to-go" function. This function measures the expected future cost of each state of a system and, if known, can be used to pick the best course of action for the control, much as a chess player chooses a move that will create or avoid checkmate. (Chess, however, is a game of perfect information and a finite number of moves; stochastic control problems are "games" of chance that typically go on forever.) The "cost" of a game is defined as $c(s_0, u_0) + \beta c(s_1, u_1) + \beta^2 c(s_2, u_2) + \dots$, where $s_0, s_1, s_2, \dots$ is the sequence of states, $u_0, u_1, u_2, \dots$ are the control actions, $c(s, u)$ is the cost of applying control u to state s , and $\beta < 1$ is a "discount" factor that converts future costs to their present value. Mathematically, the effect of a control action is to influence the probability with which state s_t changes to state s_{t+1} . The cost-to-go function is the expected cost of a game, starting at state s_0 , minimized over all possible control strategies, a strategy being of the form "if you're in state s , then apply control $u(s)$."

The trouble, of course, is that in applications of any consequence the cost-to-go function is rarely

known. The definition is of little help, mainly because it's based on what you're really after, namely the optimal (expected-cost-minimizing) control strategy. In many problems, however, it's possible to identify states whose cost-to-go values are obviously close. That's where feature extraction and approximations come in.

The basic idea is to settle on a relatively small number of features that the

states of a system can exhibit, and then assign an approximate cost-to-go value to each combination of features. If the number of combinations is relatively small, this can be done directly with a look-up table; alternatively, a simple mathematical formula, such as a weighted linear combination of feature values, can be used. What's important is that the approximate cost-to-go function can be adjusted during a training period. That is,

as the actual costs of applying control actions become apparent, you can change your mind about what to do when you see particular combinations of features.

The trick is to make changes that ultimately converge on a good strategy.

To be sure, any feature-based approach is limited by the "resolution" of its features.

To be sure, any feature-based approach is limited by the "resolution" of its features: If two states have the same combinations of features but their true cost-to-go values differ by, say, \$20, then no algorithm can guarantee an error any smaller. But it should be possible to get close to the resolution limit.

Not every adjustment scheme does so. In particular, adjustments to an algorithm's approximate cost-to-go function based on a least-squares fit to experience can go badly astray. Extending an example first described by Bertsekas, Tsitsiklis and Van Roy have produced an extremely simple, two-state system with a true cost-to-go function that is identically zero, but with an approximate cost-to-go function that goes to infinity when naive but natural least-squares adjustments are applied.

Tsitsiklis and Van Roy propose a different approach that provably converges to within a factor of $1/(1 - \beta)$ of the resolution limit. (The dependence on the discount factor makes sense, since the

resolution limit affects each term in the cost series $c + \beta c + \beta^2 c + \dots$) Their approach is quite simple. In essence, they decide in advance how often to update the cost-to-go approximation and how

much "weight" to give each round of new data. If the current cost-to-go approximation is C and the data of round j say that C^j would be better, their approach actually changes the approximation to $(1 - w_j)C + w_j C^j$, where $0 < w_j < 1$ is the pre-assigned weight for round j . Their theorem says that if the sum of the weights is infinite but the sum of their squares is finite, then, with probability 1, the adjustments converge to a cost-to-go approximation whose error is bounded by the product of $1/(1 - \beta)$ and the resolution limit.

The true test of an approach to any problem in applied mathematics, of course, is in actual applications. In a series of initial tests, Tsitsiklis and Van Roy have pitted their approach against a problem that people—*young people*, to be precise—excel at solving: the game of Tetris.

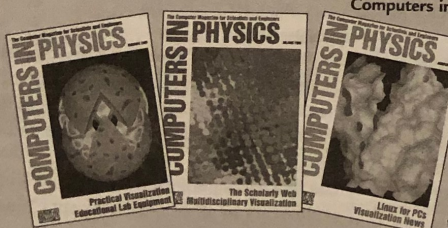
Tetris is a computer game in which the player maneuvers a random sequence of falling bricks of various shapes in order to create solid brick walls. As the game

See Control on page 13

Send for your FREE Issue of Computers in Physics.

You'll discover:

- Internet sites, tools, and resources — including a new Internet Goldmine in every issue.
- A guide to current, practical applications in research and education spelled out in Laboratory Applications, Numerical Algorithms, Computer Simulations and Computers in Education.
- Hands-on, proven techniques to save you time and money — see our Scientific Programming and Practical Visualization departments as well as Product Reviews.
- A peer-reviewed journal section — leading researchers and educators cover significant advances in computational science and education.
- And much more — features, news, commentary and interviews with people like Ken Wilson, Paul Ginsparg and Donald Knuth.



Computers in Physics is the computer magazine for scientists and engineers. Find out why with no obligation or risk. Order your FREE trial copy today!

Act now and get the latest issue of Computers in Physics absolutely FREE.

COMPUTERS IN PHYSICS FREE ISSUE OFFER

- ☒ Yes, send my FREE trial copy of Computers in Physics without obligation or commitment. If I choose to continue the subscription, I will pay the rate indicated below for an annual subscription and receive an additional 5 issues. If not, I'll write "cancel" on your invoice and owe nothing. The free issue is mine to keep.

- ☐ Individuals \$59.95 (Members of AIP Member & Affiliated Societies \$49.95)
☐ Institutions \$275

IMPORTANT: These rates are for subscribers in the U.S. For subscriptions to Canada, Mexico, Central and South America, and the Caribbean add \$10/year for surface freight. For subscriptions to Europe, Asia, The Middle East, Africa and Oceania add \$20/year for air freight.

Name _____
 Title _____
 Company _____
 Address _____
 City/State/Zip _____ SBD608

AMERICAN INSTITUTE OF PHYSICS

Mail to:
 American Institute of Physics,
 Marketing Dept., 500 Sunnyvale Blvd,
 Woodbury, NY 11797, USA

For faster service, e-mail your request to mktg@aip.org (type SBD608 in the subject line).

Control

continued from

continues, the with gaps in the player manager disappears drop down a particular column stage as the negative eliminated by positive "ret" each row the ber of points fast the play.

In devel Tsitsiklis and two crude f current wall. Using a qua features, the algorithm. T nated an ave a trial run i respectable, score. (It's l do.) When t training sess algorithm u rows per gar the algorithm pletely rand

The Dep Mathem

The pro require

• Stoch (Prob: • Num (Com • Finan (Port

The dir

The pr be take offered practit

Applic ics, PT

Appli

Office Gradu Colum 107 L New (212) ref8@

Further www

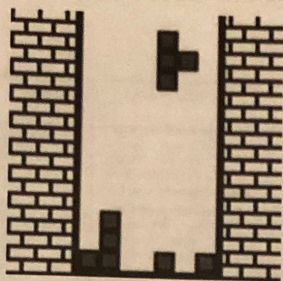
The ap

Control

continued from page 12

continues, the wall rises row by row, but with gaps in the brickwork. Each time a player manages to create a solid row, that row disappears and the bricks above it drop down a level. The "cost" of any particular control action taken at a particular stage of the game can be defined as the negative of the number of rows eliminated by that action. (In actual play, positive "reward" points are given for each row that disappears, with the number of points depending in part on how fast the play is proceeding at that point.)

In developing their algorithms, Tsitsiklis and Van Roy chose to focus on two crude features: the height of the current wall and the number of gaps in it. Using a quadratic function of the two features, they then wrote a rule-based algorithm. This "expert player" eliminated an average of 31 rows per game in a trial run of 100 games, which is a respectable, though not outstanding, score. (It's better than most adults can do.) When they used these games as a training session, the "new and improved" algorithm upped its score to about 32 rows per game. On the other hand, when the algorithm was "trained" on completely random Tetris configurations, the



After tests on the computer game Tetris—in some sense a worst-case scenario for testing automatic control algorithms—John Tsitsiklis and Dimitri Bertsekas see a bright future for feature-based algorithms that learn.

score fell to about 11 rows per game.

Tsitsiklis describes the admittedly unimpressive performance of their initial algorithm as a "sanity check" that the general approach could do something sensible, even if it's not given much to work with. In particular, the poor play of the randomly trained algorithm indicates that the sampling distribution for the training set plays a crucial role in the performance of feature-based algorithms that learn from experience. (As any math teacher can tell you, you'll do better on your trig exam if you bone up on sines

and cosines rather than logarithms. The result was still well within the theoretical bound. Tsitsiklis and Van Roy used a discount factor β of 0.9999.)

A smarter, more observant program—one that takes into account more features and uses a more complex learning algorithm—could presumably do better. And indeed, that's precisely what's happened. Bertsekas and Sergey Ioffe, a graduate student at MIT, have trained a Tetris-playing program that can eliminate more than 3000 rows in an average game. That's better, even, than what most kids can do. (Moreover, playing Tetris doesn't turn the computer into a zombie.)

"In some sense, Tetris is a worst-case scenario for the evaluation of automatic control algorithms, since humans excel at Tetris," Tsitsiklis says. But given that an algorithm based on two crude features can do respectably, and that a more sophisticated version can do far better, he sees a bright future for feature-based algorithms that learn. Even so, intelligent control systems may, by definition, never catch up with human beings. The reason is simple, Tsitsiklis says: "We call 'intelligent' anything which is one step ahead of what we know how to do well."

Barry A. Cipra is a mathematician and writer based in Northfield, Minnesota.

Fellows

continued from page 16

used at Los Alamos in considering candidates for laboratory fellow; one of the items on the list is whether the candidate is a fellow of his or her primary professional society. Promotion to high-level administrative positions can also hinge on honors that are not recognized by the mathematical sciences community. The lack of a fellow designation within the mathematical sciences community could be handicapping SIAM members.

SIAM as it currently exists is "a model of a good society," says former vice president for education Gil Strang of MIT. "Its meetings and publications have real content. Its dues are reasonable. In everything I have seen, its leadership is unselfish. I would like our society to grow in influence, because that influence would be for good."

Strang opposes the creation of a SIAM fellows program, however. "One great attraction of applied mathematics is that everybody contributes," he says. "We don't need gods and we don't have them. Key ideas like the FFT and SVD will be forever anonymous. Writing good software is a way to help science go forward, not a way to be famous. SIAM has avoided elitism so successfully that I feel sure we will not start down this road."

A better way to recognize achievement, Strang believes, is to create multiple awards for good papers and good software—"these are recognitions that have meaning for professionals."

SIAM currently recognizes outstanding achievement through its prizes, which, while prestigious, are awarded to relatively few people. The selection of invited speakers for SIAM meetings is a mechanism that recognizes larger numbers of people. One new program that has been suggested is a "paper of the year" for each of the SIAM journals, chosen by vote of the individual editorial boards. With the exception of the SIAM Prize for Distinguished Service to the Profession, however, none of these existing or proposed programs take into account contributions to the organization or service to the community.

There is clearly a diversity of opinions and feelings on the issue of rewards in general, and the question of a SIAM fellows program in particular. The board encourages SIAM members to complete the Web questionnaire on the subject (<http://www.siam.org/fellows.htm>); alternatively, members can express their opinions and suggest new programs via e-mail (littleton@siam.org). In either case, responses should reach SIAM no later than July 1.—JMC

Columbia University

Department of Mathematics

Master of Arts in Mathematics

With Specialization in the Mathematics of Finance

Fall 1997

The Department of Mathematics at Columbia University invites applications for a new track of its "Master of Arts in Mathematics" program, with specialization in the Mathematics of Finance.

The program is co-sponsored with the Department of Statistics. It concentrates on the advanced quantitative methods required for modern finance and draws on the diverse strengths of Columbia University in

- **Stochastics**
(Probability, Statistics, Random Processes)
- **Numerics**
(Computational Methods, Partial Differential Equations), and
- **Finance Applications**
(Portfolio Analysis, Option Pricing, Interest-Rate Modeling).

The director of the program is Ioannis Karatzas, Higgins Professor of Applied Probability at Columbia University.

The program is scheduled to start in September 1997, pending approval by the New York State Board of Regents. It can be taken on either a full-time (one year) or a part-time (three or more semesters) basis. All the required courses will be offered in the evenings. An important component of the program will be its weekly seminar, in which distinguished practitioners will be invited to discuss their work, give mini-courses, and suggest open problems.

Prerequisites

Applicants should have a very substantial background in Mathematics at the level of an undergraduate major in Mathematics, Physics, or Science/Engineering.

Application Procedure

Application forms for fall 1997 are available from:

Office of Graduate Admissions
Graduate School of Arts and Sciences
Columbia University
107 Low Memorial Library, Mail Code 4303
New York, NY 10027
(212) 854-4737
ref8@columbia.edu

Further information and a downloadable application package are available at the department's Web site, http://www.math.columbia.edu/departments/masters_finance.html.

The application deadline is **June 1, 1997**, but late applications will be considered.

Florida A&M Update

Due to an unfortunate miscommunication, the *SIAM News* article "IMA Workshop Highlights Contributions of Minorities to Industrial Mathematics" (December 1996) failed to reflect the fact that when James Turner joined the faculty of Florida A&M University in Tallahassee in 1995, he was not the first black PhD at that institution. He was actually following in the footsteps of such African-American PhDs and pioneers in the mathematical sciences as Israel Glover, J. Arthur Jones, Osiefield Anderson, Emma Rose Fenceroy, Roselyn Williams, Henry L. Williams, Donald M. Cole, and James E. Glover.