PERGAMON

Automatica 35 (1999) 331–338

Brief Paper

An ordinal optimization approach to optimal control problems[1]

Mei (May) Deng[a,*], Yu-Chi Ho[b]

[a] AT&T Labs, Room 1L-208, 101 Crawfords Corner Road, Holmdel, NJ 07733, USA
[b] Division of Applied Science, Harvard University, 29 Oxford Street, Cambridge, MA 02138, USA

Received 6 November 1995; revised 13 January 1997; revised in final form 24 March 1998

Abstract

We introduce an ordinal optimization approach to the study of optimal control law design. As illustration of the methodology, we find the optimal feedback control law for a simple LQG problem without the benefit of theory. For the famous unsolved Witsenhausen problem (1968, SIAM J. Control, 6(1)), a solution that is 50% better than the Witsenhausen solution is found. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Artificial intelligence; Optimal control; Optimization; Search methods; Stochastic systems; Simulation

1. Introduction

Ordinal optimization (OO) is a method of speeding up the process of stochastic optimization via parametric simulation (Deng et al., 1992; Ho, 1994; Ho and Larson, 1995; Ho and Deng, 1994; Ho et al., 1992; Lau and Ho, 1997). The main idea of OO is based on two tenets: (i) IT IS MUCH EASIER TO DETERMINE "ORDER" THAN "VALUE". This is intuitively reasonable. To determine whether A is greater or less than B is a simpler task than to determine the value of $A-B$ in stochastic situations. Recent results actually quantified this advantage (Dai, 1997; Lau and Ho, 1997; Xie, 1997). (ii) SOFTENING THE GOAL OF OPTIMIZATION ALSO MAKES THE PROBLEM EASIER. Instead of asking the "best for sure" we settle for the "good enough with high probability". For example, consider a search on design space Θ. We can define the "good enough" subset, $G \subset \Theta$, as the top-1% of the design space based on system performances, and the "selected" subset, $S \subset \Theta$, as the estimated (however approximately) top-1% of the design choices. By requiring the probability of $|G \cap S| \neq 0$ to be very high, we insure that by narrowing the search from Θ to S we are not "throwing out the baby with the bath water". This again has been quantitatively reported in Deng (1995), Lau and Ho (1997) and Lee et al. (1998).

Many examples of the use of OO to speed up the simulation/optimization processes by orders of magnitude in computation have been demonstrated in the past few years (Ganz and Wang, 1994; Ho and Larson, 1995; Ho and Deng, 1994; Ho et al., 1992; Lau and Ho, 1997; Patsis et al., 1997; Wieseltheir et al., 1995). However, OO still has limitations as it stands. One key drawback is the fact that Θ for many problems can be HUGE due to combinatorial explosion. Suppose $\Theta = 10^{10}$ which is small by combinatorial standards. To be able to get within the top-1% of Θ in "order" is still 10^8 away from the optimum. This is often of scant comfort to optimizers. The purpose of this note is to address this limitation through iterative use of OO very much in the spirit of hill climbing in traditional optimization.

2. Model and concepts

Consider the expected performance function $J(\theta) = E[L(x(t;\theta,\xi))] = E[L(\theta,\xi)]$, where $L(x(t;\theta,\xi))$ represents some sample performance function evaluated through the realization of a system trajectory $x(t;\theta,\xi)$ under the design parameter θ. Here ξ represents all the

* Corresponding author. Tel.: 001 732 949 7624; fax: 001 732 949 1720; e-mail: mdeng@att.com.

[1] This paper was not presented at any IFAC meeting. This paper was recommended for publication by Associate Editor Arunabha Bagchi under the direction of Editor Tamer Başar. The work reported in this paper is partially supported by NSF grants EEC-9402384, EEC-9507422, Air Force contract F49620-95-1-0131, and Army contracts DAAL-04-95-1-0148, DAAL-03-92-G-0115.

PII: S0005-1098(98)00155-1

random effects of the system. Denote by Θ, a huge but finite set, the set of all admissible design parameters. Without loss of generality, we consider the optimization problem $\mathrm{Min}_{\theta \in \Theta} J(\theta)$. In OO, we are concerned with those problems where $J(\theta)$ has little analytical structure but large uncertainty and must be estimated through repeated simulation of sample performances, i.e., $J(\theta)_{\mathrm{est}} = (1/K)\sum_{i=1}^{K} L(\theta, \xi_i)$, where ξ_i is the ith sample realization of system trajectory or often equivalently, $J(\theta)_{\mathrm{est}} = L(x(t; \theta, \xi))$; $t -> \infty$. The principal claim of OO is that performance **order** is relatively robust with respect to very small K or $t \ll \infty$ (we shall use short simulation or small number of replications interchangeably here after to indicate that the confidence interval of the performance value estimate is very large). More specifically, let us thus postulate that we observe or estimate $\tilde{J}(\theta) = J(\theta) + w(\theta)$, where $w(\theta)$ is the estimation error or noise associated with the observation/estimation of design θ. We assume that $w(\theta)$ is a 0-mean random variable which may have very large variance (corresponding to the fact that $\tilde{J}(\theta)$ is estimated very approximately which simplifies the computational burden). In most of our experiments, the performance $J(\theta)$ is estimated by statistical simulation. We observe that $w(\theta)$'s depend mainly on the length (number of replications) of the simulation experiments and are not very sensitive to particular designs. Furthermore, as demonstrated by Deng et al. (1992), the methodology of OO is quite immune to correlation between noises. In fact, correlation in estimation errors in general helps rather than hinders the OO methodology. When the performance $J(\theta)$ is basically deterministic but very complex and computationally intensive and a crude $\hat{J}(\theta)$ is used to approximate it, then $w(\theta)$'s cannot generally be considered as i.i.d. since $w(\theta)$'s are the results of approximation errors rather than statistical estimation errors. In such a case, we have to uncorrelate the data before the OO theory can be applied. The details of this approach is described in Lee et al. (1997). In this paper, we assume that $w(\theta)$'s are design-independent.

We next introduce a thought experiment of evaluating $J(\theta)$ for all $\theta \in \Theta$ and plot the histogram of the distribution of these performances. With slight abuse of terminology (since Θ is huge, the histogram can essentially be considered as continuous and as a probability density function), we call this histogram the performance density function and its integral the performance distribution function (PDF). If we now uniformly take N samples of $J(\theta)$, then it is well known that the representativeness of the sample distribution to the underlying PDF is independent of the size of the population or $|\Theta|$ and only depends on the sample size N. In fact, using Chebyshev's inequality, we have for any $\varepsilon > 0$, $\mathrm{Prob}[|\text{sample PDF} - \text{PDF}| \geq \varepsilon] \leq 1/4N\varepsilon^2$. For example, $N = 1000$ can guarantee that with probability ≥ 0.975 the sample PDF is close to the true PDF within 0.1 and $N = 5000$ will make $\varepsilon = 0.045$ for the same probability. Now for our problem we can only take noisy samples of $J(\theta)$, i.e. $\tilde{J}(\theta)$. Yet it is still true that the accuracy of any estimate about the PDF of $J(\theta)$ based on these noisy samples is only a function of the size of samples and the magnitude of the noise and is independent on the size of the underlying population.

3. General idea and methodology

A major difficulty of many stochastic optimization problems comes from the combinatorial explosion of the search spaces associated with the problems. For example, in an optimal control problem of a continuous variable dynamic system, the state space is usually a continuous set. A control rule is a mapping from the state variables to control variables. Therefore, the space of control rule designs even with the discretization of state and control variables is a space that is essentially infinite and does not have a simple structure. When the search space of a problem is large and complicated, little analytical information is available, and the estimation of performance is time consuming, it is practically impossible to find the optimal solution. In this note, we propose and demonstrate an iterative and probabilistic search method whose aim is to find some solutions which are not necessary for the optima but "good enough".

To help fix ideas, consider the case where 5000 alternatives in a space of $\Theta = 10^{10}$ are uniformly sampled. What is the probability that among the samples at least one belongs in the top-k designs of the space Θ? The answer can be very easily calculated, Prob(at least one of 5000 samples is in top-k) $= 1 - (1 - k/10^{10})^{5000}$, which for $k = 50$, 500, 5000, are 0.00002, 0.00025, and 0.0025, respectively. Not at all encouraging! But if now somehow we "**know**" that the top-k designs are concentrated in a subset of 10^6, then the same number of 5000 samples can *improve* the probability to 0.222, 0.918, and certainty, respectively, i.e., we can hardly fail! While these are two extreme cases, we can interpolate between these cases where our "knowledge" is less crisp and more uncertain. The point is the importance of knowledge. In soft computing or computing intelligence (Jang et al., 1997), the same issue is referred to as "encoding" as in genetic algorithms, or knowledge "representation" or "bias selection" in AI (Ruml et al., 1997). In other words we should bias or narrow our search to favor a subset of the search using knowledge. How knowledge is acquired is still an open problem in AI (Jang et al., 1997, p. 434). However, heuristically, one can learn from experience. In particular, if we sample a set of designs for their performances (however, noisily or approximately) then we should be able to gleam from the samples what are "good" subspaces to search and gradually restrict the search there. This is in the spirit of traditional hill climbing except instead of moving from point to point in the search space, we move from one subset or search representation to another. The speed up of OO enables us to do such iterations on search subspaces.

The key here is to establish a procedure of comparing two representations of a search space based upon sampling and then to narrow the search space step by step. How do we compare two representations of a search space? Suppose Θ_1 and Θ_2 are two subspaces of a large search space Θ. Does Θ_1 have more good designs than Θ_2? It seems that in order to answer the question we must fully investigate Θ_1 and Θ_2 which may be time-consuming and even practically impossible when the spaces are large and there exist large observation noises. However, using our OO methodology, we may answer the question with a high confidence level by simply comparing the observed PDFs of Θ_1 and Θ_2.

Lemma. *Suppose $J(\theta)$ is a measurable function and $w(\theta)$'s are design-independent continuous random variables. For $\tilde{J}(\theta) = J(\theta) + w(\theta)$, if we independently sample two designs θ_1 and θ_2 according to a common sampling rule from the search space Θ and observe that $\tilde{J}(\theta_1) < \tilde{J}(\theta_2)$, then* $\mathrm{Prob}(J(\theta_1) \le J(\theta_2) | \tilde{J}(\theta_1) < \tilde{J}(\theta_2)) \ge 0.5$ (*Contact the authors for the proof.*)

The importance of the Lemma is that "seeing is believing". We should believe in what we see even in the presence of large noises. We can extend our observation of "seeing is believing" to compare two representations of a search space using their observed PDFs. Suppose Θ_1 and Θ_2 are two representations (subspaces) of a large search space Θ. Let $F_1(t)$ and $F_2(t)$ be the observed PDFs of Θ_1 and Θ_2, respectively. If $F_1(t) \ge F_2(t)$ for $t < t_1$ where t_1 is determined by the satisfaction level (e.g. to search top 5%, designs, t_1 is the value such that $F_1(t_1) = 0.05$), then Θ_1 is more likely to have good designs than Θ_2 and we should continue our search in Θ_1. More Details on space comparison can be found in Deng (1995).

Now we are ready to summarize our sampling and space-narrowing procedure. For a search space Θ, we first define two or more representations (subsets of searches) and find their corresponding observed PDFs. By comparing the observed PDFs, we can identify which representation(s) is (or are) good probabilistically. We can then further narrow our search into smaller subspaces. The above process is a man–machine interaction iteration. In the next section, we demonstrate the sampling and space-narrowing procedure through applications.

4. Applications

We study two well-known continuous time optimal control problems. The first one is the famous Witsenhausen problem (Witsenhausen, 1968). It is an extremely simple Linear–Quadratic-Gaussian (LQG) problem to state but exceedingly hard to solve. The twist is the presence of nonclassical information pattern. No optimal solution has yet been found for the Witsenhausen problem (WP). Witsenhausen has presented a nonoptimal nonlinear solution that is at least 50% better in performance than the best linear solution. The second problem is a simple LQG problem, that is, a linear control problem with quadratic performance function and additive Gaussian white noises. It is known that the optimal solution of the problem is a linear control law. For both problems, the design space can be reduced to find a set of mappings from a one-dimensional state space to a one-dimensional control space. If we discretize the two variables x (state) and u (control) to n and m values each, the set of all admissible control laws, Γ (the space Θ), has the size of m^n.

4.1. The Witsenhausen problem

In this subsection, we apply our sampling and space-narrowing procedure to study WP which is well known and of long standing. It is an extremely simple scalar two-stage LQG problem with the twist that the information structure is nonclassical, i.e. the control at the second stage does not remember what it knows at the first stage. WP presents a remarkable counterexample which shows that the optimal control laws of LQG problems may not always be linear when there is imperfect memory. The optimal control law of WP is still unknown after 28 years. The discrete version of the problem is known to be NP-complete (Papadimitriou and Tsitsiklis, 1986).

The problem is described as follows. At stage 0 we observe z_0 which is just the initial state x_0. Then we choose a control $u_1 = \gamma_1(z_0)$ and the new state will be $x_1 = x_0 + u_1$. At stage 1, we can not observe x_1 directly, instead, we can only observe $z_1 = x_1 + v$ where v is a noise. Then we choose a control $u_2 = \gamma_2(z_1)$ and system stops at $x_2 = x_1 - u_2$. The cost function is $E[k^2(u_1)^2 + (x_2)^2]$ with $k^2 > 0$ a constant. The problem is to find a pair of control functions (γ_1, γ_2) which minimizes the cost function. The trade off is between the costly control of γ_1 which has perfect information and the costless control γ_2 which has noisy information. First, we consider the case when $x_0 \sim \mathrm{N}(0, \sigma^2)$ and $v \sim \mathrm{N}(0, 1)$ with $\sigma = 5$ and $k = 0.2$.

Witsenhausen made a transformation from (γ_1, γ_2) to (f, g) where $f(z_0) = z_0 + \gamma_1(z_0)$ and $g(z_1) = \gamma_2(z_1)$. Then, the problem is to find a pair of functions (f, g) to minimize $J(f, g)$ where $J(f, g) = E[k^2(f(x_0) - x_0)^2 + (f(x_0) - g(f(x_0) + v))^2]$. Witsenhausen (Witsenhausen, 1968) proved the following: 1. For any $k^2 > 0$, the problem has an optimal solution. 2. For any $k^2 < 0.25$ and $\sigma = k^{-1}$, the optimal solution in linear controller class with $f(x) = \lambda x$ and $g(y) = \mu y$ has $J^*_{\mathrm{lin}} = 1 - k^2$, and

$\lambda = \mu = 0.5\,(1 + \sqrt{1 - 4k^2})$. When $k = 0.2$, $J^*_{\text{lin}} = 0.96$. 3. There exist k and σ such that J^*, the optimal cost, is less than J^*_{lin}, the optimal cost achievable in the class of linear controls. Witsenhausen gave the following example. Consider the design: $u_1 = -z_0 + \sigma \text{sgn}(z_0)$, $u_2 = \sigma \tanh(\sigma z_1)$, the cost function J is $J_{\text{wit}} = 0.4042532$. 4. The optimal control law (f^*, g^*) is still not known. But given the function f, the optimal g^*_f associated with function f is $g^*_f = E[f(x_0)\,\varphi(z_1 - f(x_0))]/E[\varphi(z_1 - f(x_0))]$.

Now the problem becomes to search for a single function f to minimize $J(f, g^*_f)$. Although the problem looks simple, no analytical method is available yet to determine the optimal f^*. However, there are some properties of the optimal control function f^*: $E[f^*(x)] = 0$ and $E[(f^*(x))^2] \le 4\sigma^2$.

Next, we demonstrate how to apply our sampling and space-narrowing procedure to search for good control laws for WP. The controllers γ_1 and γ_2 are constructed as follows: 1. Based on the property $E[f^*(x)] = 0$, we make the assumption that γ_1 is symmetric, i.e., $\gamma_1(z_0) = -\gamma_1(-z_0)$. 2. The function $f(z_0) = \gamma_1(z_0) + z_0$ is a staircase function constructed by the following procedure. (1). Divide the z_0-space $[0, \infty)$ into n intervals, $I_1, \ldots, I_n$, where $I_i = [\sigma t_{(0.5+0.5(i-1)/n)}, \sigma t_{(0.5+0.5*i/n)})$. t_α is defined by $\Phi(t_\alpha) = \alpha$ where Φ is the standard normal distribution function. $\text{Prob}[z_0 \in I_i] = 0.5/n$ because z_0 has a normal distribution $N(0, \sigma^2)$. This implies that we discretize the z_0-space evenly in probability. (2). For each interval I_i, a control value f_i is uniformly picked from $(-3\sigma, 3\sigma)$, i.e. $f_i \sim U(-15, 15)$. 3. For any function f constructed in step 2, the controller $\gamma_2(z_1) = g^*_f(z_1)$ is computed based on an approximation of g^*_f, $\hat{g}_f = \sum_{i=1}^{100} f(x_i)\varphi(z_1 - f(x_i)) / \sum_{i=1}^{100} \varphi(z_1 - f(x_i))$.

The search is conducted using the following procedure. We construct the control function f for u_1. Then we compute control function $\hat{g}_f$ and get the performance $\tilde{J}$ of the two-stage problem by 100 replications. Here the observed performance $\hat{J}$ is obtained with relatively large noise because we only run 100 replications. Five thousand pairs of controllers $(f, \hat{g}_f)$ are generated to construct the estimated PDF.

Four values of $n = 1$, 2, 5, and 10, the number of intervals in the construction of f, are selected first. Fig. 1 shows the observed PDFs. From sampled PDFs, we observe that the cases of $n = 5$ and 10 are worse than those of $n = 1$ and 2. The cases of $n = 1$ and 2 are indistinguishable at this stage. However, we can see that the space with two intervals is the best among the four when the satisfaction level is small. The fact that $n = 1$ is a good choice indicates that the class of constant control functions (with discontinuity at the origin due to symmetry) is a good representation of the search space for control function f. This means that the pair of controllers described by Witsenhausen which outperforms the optimal linear control law is already very good. From the observed best control functions for different n, we observe that the control values of the observed best controllers for different n are located in $[-2, 12)$. Based on this observation, we make the following restriction (W1): *For each interval I_i, control f is in $(-0.5\sigma, 2.5\sigma)$, i.e. $f \sim U[-2.5, 12.5)$.*

With this restriction, we repeat our experiment for $n = 1, 2, 5$, and 10. We observe the same phenomenon as in Fig. 1. This encourages us to search in the spaces of 3 and 4 intervals with restriction W1. Figs. 2 and 3 show the PDFs and observed best-control functions for 1, 2, 3 and 4 intervals. From Fig. 2, we may conclude that the space with two intervals is the best representation we have found at this stage. A more interesting phenomenon we may observe from Fig. 3 is that the observed best

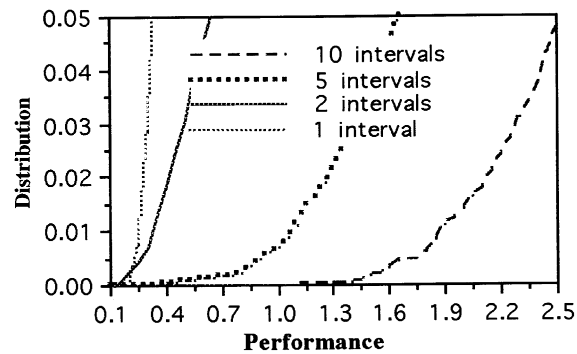

Fig. 1. Observed PDFs with 1, 2, 5 and 10 intervals.

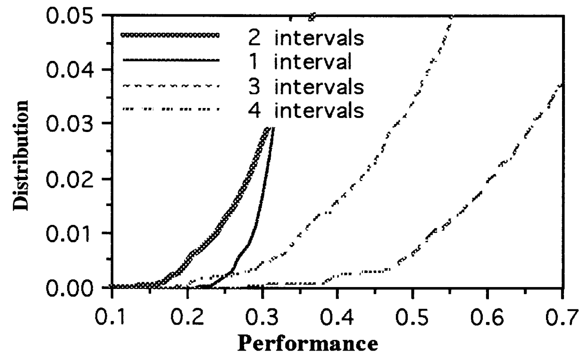

Fig. 2. Observed PDFs with restriction W1.

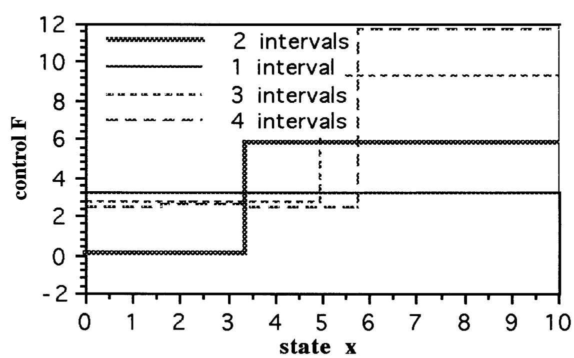

Fig. 3. Observed best control rules with restriction W1.

controllers for both $n = 3$ and 4 have the two-interval shape as the one of $n = 2$. This further indicates that the right direction of search may be toward the two-interval functions. Since the observed best controllers display some increasing property, we make a further restriction (W2): *The control f is an increasing function in* $(-0.5\sigma, 2.5\sigma)$.

In Fig. 4, the curves without the legend " interval (in)" are those without W2. Fig. 4 shows that with restriction W2, the two-interval controllers have the best PDF. This indicates that the specification of increasing control function is a right direction. However, the space of three-interval is worse than the space of two-interval and the space of four-interval is worse than the space of three-interval. Furthermore, we observe the same phenomenon as in Fig. 3, the observed best controllers of $n = 3$ and 4 have the two-interval shape. It is conceivable that the optimal control function may possess significant discontinuity. So the 3rd restriction (W3) is made as follows: *The control f is a two-value increasing step function in* $(-0.5\sigma, 2.5\sigma)$.

Next, we want to determine the jump point of the two-value functions. The jump point of the two-value functions we have examined so far is at $\sigma t_{0.75}$. For simplicity, we only check the jump points $\sigma t_{0.55}$, $\sigma t_{0.60}, \ldots, \sigma t_{0.95}$. The PDFs associated with different jump points are presented in Fig. 5. In Fig. 5, the "2 int. (in)" represents the two-interval increasing controllers we used before which has the jump point at $\sigma t_{0.75}$. We see that the best jump point is around $\sigma t_{0.90}$. The best observed control function f among 5000 samples in the space associated with $\sigma t_{0.90}$ jump point is $f(x) = 3.1686$, $0 \le x < 6.41$; and $f(x) = 9.0479$, $x \ge 6.41$. The estimated true performance (obtained by 10 000 replications) for the best observed control function f is 0.190 with variance 0.0001.

We also have conducted the experiment to search the optimal polynomial control function f with the symmetricity restriction. The best control function we found is a symmetric function, $f(x) = 0.0228x + 3.8019$ for $x > 0$. And its estimated true performance is 0.3541 with variance 0.00005. Although this is a better solution than Witsenhausen's example, it cannot compete with the ob-

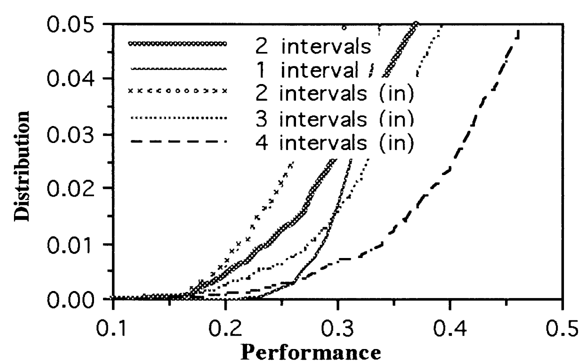

Fig. 4. Observed PDFs with restriction W2.

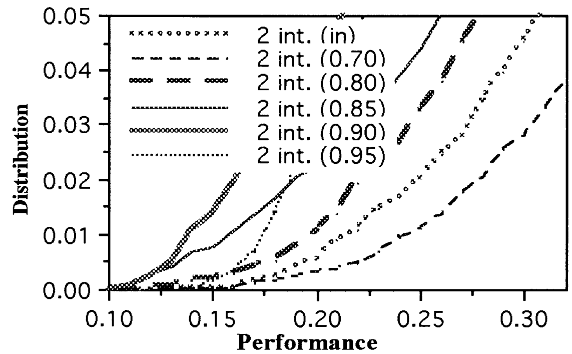

Fig. 5. Observed PDFs with restriction W3.

served best control function found in the two-value control function space with $\sigma t_{0.90}$ jump point.

For benchmarking, we choose the case when $x_0 \sim N(0, 100)$, $v \sim N(0, 1)$ and $k^2 = 0.1$. We use the same approach as described above to find a good control function f in the space with two-value control functions. Our result J_{oo} and other results reported in Baglietto et al. (1997) are listed below. The best linear solution $J_{lin} = 0.90$. The Witsenhausen nonlinear solution $J_{wit} = 0.428$. The optimized Witsenhausen solution (Bansal and Basar, 1987) $J_{Wopt} = 0.417$. The neural network solution (Baglietto et al., 1997) $J_{nn} = 0.409$. The iterative OO solution $J_{oo} = 0.334$. The control function associated with J_{oo} is $f(x) = 0.1024$, $0 \le x < 2.66$; and $f(x) = 5.4450$, $x \ge 2.66$.

The after-the-fact reasoning behind the superiority of the two-value controllers is as follows. Instead of the original strategy of converting x_1 to a two point signaling distribution (one step function), we convert x_1 to a four point signaling distribution strategy (two step function). For the parameters used in this paper, the error committed in the four point signaling strategy is very small. At the same time, the four point strategy saves more control energy cost.

4.2. A simple LQG problem

Consider the following well-known LQG problem. The stochastic system is represented by $x_{t+1} = -x_t + u_t + w_t$; $t = 0, \ldots,$ where $w_t \sim N(0, 25)$ and the performance criterion is $J = E\{\lim_{T\to\infty}(1/T)\sum_{i=1}^{T} x_i^2 + u_i^2\}$. We wish to determine the optimal feedback control law $u_t = \gamma(x_t)$ among all $\gamma \in \Gamma$. *First*, it is easy to argue that the optimal control law is symmetric. *Second*, because the performance function is the sum of squares of states and controls, to minimize the performance and stabilize the system, the controls should keep the states around 0. Since the variance is 25, it is reasonable to believe that both states and controls will hardly go beyond $[-100, 100]$. So we make the following restriction: If $0 \le x_t < 100$, $u_t \in [0, 100)$ and if $x_t \ge 100$, $u_t = 100$. Next, we digitize the state interval $[0, 100)$ into n intervals equally. Note that, according to the assumptions, we have another n intervals in $[-100, 0)$ and two more intervals for $x \geq 100$ and $x < -100$, respectively. Let us denote the discretized search space with $2n+2$ intervals as Γ^n. Now we can define a feedback control law in this discretized space as follows: *For the ith* $(i = 1, \ldots, n)$ *interval in* $[0, 100)$, *we randomly choose a value for* u_t, $u_t \sim U(0, 100)$, *i.e.*, $m -> \infty$. *The states in one interval will share the same control value generated for that interval, i.e. the control law is a step function.*

In our experiment, we sample 5000 control laws uniformly in Γ^n. For each control law we obtain its observed performance $\hat{J}$ by very short simulation in the sense that the time horizon is $T = 100$ (instead of $T -> \infty$). We cannot overemphasize the importance of ordinal optimization which enables us to estimate performances with very crude models (i.e. ultra short simulations with large estimation errors). To evaluate 5000 control laws to steady state would be computationally burdensome if not infeasible.

We will use the similar man–machine iterative searching procedure as in Section 4.1. We first consider the cases when $n = 5$, 10, 25, 50, and 100. The observed sample PDFs of the cases of x_t-quantization are plotted in Fig. 6. Fig. 6 shows a clear pattern that the smaller the n, the better the performance. This is understandable since by using a smaller discretization of the state space, we have unwittingly captured the good property of "continuity" in feedback control laws for this problem, i.e., nearby state values should induce nearby control values. Therefore, next we consider the cases $n = 1$, 2, 3, 4, and 5. The observed PDFs are shown in Fig. 7. From Fig. 7, we see that the case of $n = 1$ is worse than others which means the constant control law is not a good one. The observed best control laws display an interesting pattern. When the states are small, the controls are small and when the states are large, the controls are large. Therefore, we make a restriction (L1): *For a given n, if* $x \in [0, 100/n)$, $u = \gamma(x) \in [0, 25)$ *and if* $x \in [(n\text{-}1)100/n, 100)$, $u = \gamma(x) \in [25, 100)$.

The generation of control laws is modified. For the first interval, $u_t \sim U(0, 25)$. For the last interval, $u_t \sim U(25, 100)$. For other intervals, $u_t \sim U(0, 100)$. Let $\Gamma^n(\text{ln})$ be the search space with restriction L1. Figs. 8 and 9 show the observed PDFs and observed best control laws for $n = 2$, 3, 4, 5. Fig. 8 verifies that restriction L1 points to the right direction of narrowing the search space. The curves with the legend "$n = 3$ (old)" and "$n = 4$ (old)" are just the curves in Fig. 7 with the legend "$n = 3$" and "$n = 4$". In Fig. 9, a clear picture emerges. All the observed best control laws in the four spaces are increasing functions. This suggests that the control law may be an increasing function. So we have the next restriction (L2): $\gamma(x)$ *is monotonically nondecreasing in x.*

Let $\Gamma^n(\text{in})$ be the search space with restriction L2. From Fig. 10, we see that restriction L2 yields better

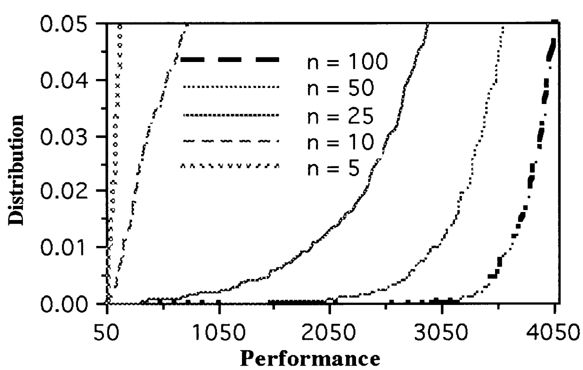

Fig. 6. Observed PDFs of Γ^5, Γ^{10}, Γ^{25}, Γ^{50} and Γ^{100}.

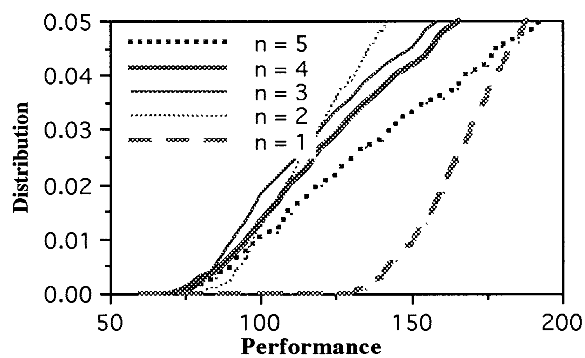

Fig. 7. Observed PDFs of Γ^1, Γ^2, Γ^3, Γ^4 and Γ^5.

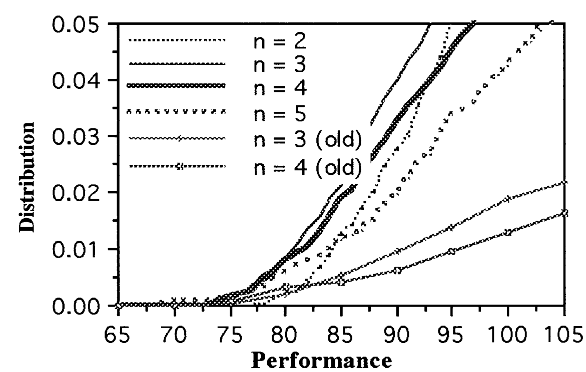

Fig. 8. Observed PDFs with restriction L1.

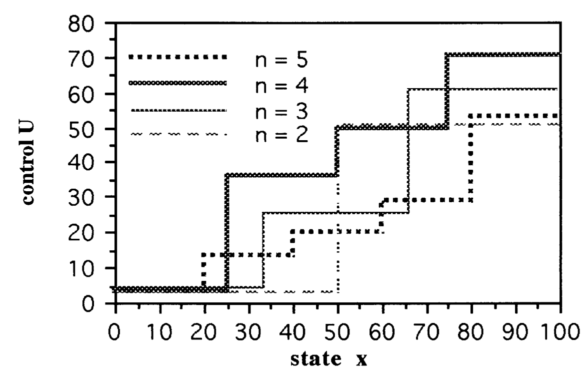

Fig. 9. Observed best control rules with restriction L1.

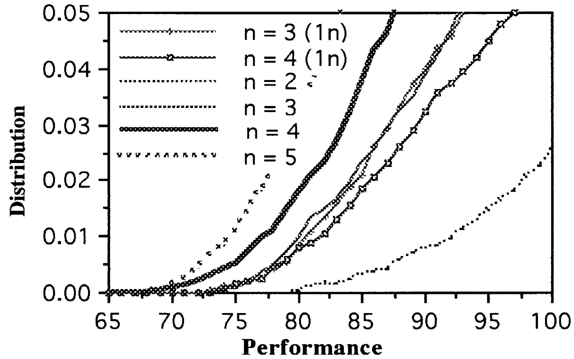

Fig. 10. Observed PDFs with restriction L2.

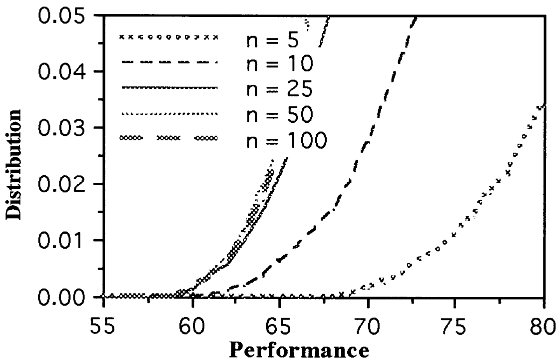

Fig. 11. Observed PDFs with restriction L2.

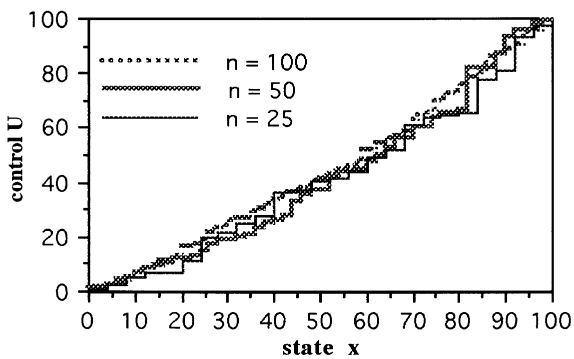

Fig. 12. Observed best control rules for Γ^{25}(in), Γ^{50}(in) and Γ^{100}(in).

results. Also from Fig. 10, we find that better spaces are now associated with larger n. So we study the search spaces associated with $n = 10$, 25, 50 and 100 under restriction L2. Fig. 11 shows the observed PDFs for $n = 5$, 10, 25, 50, 100. We see from Fig. 11 that $n = 100$ and $n = 50$ are two good representations of the search space under restriction L2. Fig. 12 shows the observed best controls for $n = 25$, 50 and 100. The performance of the observed best control law in Γ^{100}(in) is 42.19, and the performance of the observed best control law in Γ^{50}(in) is 42.33. These are to be compared with the known optimal feedback control law of $u_t = \gamma^*(x_t) = 0.618x_t$ and $J^* = 40.46$. The control laws in Fig. 12 display a pattern of linearity. We may continue this man–machine iteration process to improve the result. For example, we may establish the restriction of linearity and determine the optimal coefficient of linear control laws. But the point has been made and we shall not pursue this problem further.

5. Conclusion

How to iterate successively using OO search under the man–machine interaction paradigm is still in its infancy. There still exist a lot of open problems before this approach can be totally automated. This note, however, does illustrate the potentials of this approach in solving "hard" stochastic optimization problems with huge search space and not apparent structure. Further research along this line will also have impact on the outstanding AI problem of knowledge acquisition.

Acknowledgements

The authors like to thank Prof. LiYi Dai of Washington University and Prof. Chun-Hung Chen of University of Pennsylvania for helpful discussions.

References

Baglietto, M., Parisini, T., & Zoppoli, R. (1997). Nonlinear approximations for the solution of team optimal control problems. *Proc. 36th Conf. on Decision and Control* (pp. 4592–4594).

Bansal, R., & Basar, T. (1987). Stochastic teams with nonclassical information revisited: when is an affine law optimal?. *IEEE Trans. Automat. Control*, *32*(6), 554–559.

Dai, L. (1997). Convergence properties of ordinal comparison in the simulation of discrete event dynamic systems. *JOTA*, to appear.

Deng, M. (1995). Sampling-selection and space-narrowing methods for stochastic optimization. Ph.D. thesis, Harvard University, Harward.

Deng, M., Ho, Y.C., & Hu, J.Q. (1992). Effect of correlated estimation error in ordinal optimization. *Proc. Winter Simulation Conf.* (pp. 466–475).

Ganz, A., & Wang, X. (1994). Efficient algorithm for virtual topology design in multihop lightwave networks. *IEEE/ACM Trans. Networking, 2.*

Ho, Y.C. (1994), A new paradigm for stochastic optimization and parallel simulation. *Proc. 1993 DES Workshop in IMA/U.Minn Lecture Notes Series.* Springer, New York.

Ho, Y.C., & Deng, M. (1994). The problem of large search space in stochastic optimization. *Proc. 33rd CDC* (pp 1470–1475).

Ho, Y.C., & Larson, M. (1995). Ordinal optimization and rare event probability simulation. *J. DEDS*, *5*(2–3).

Ho, Y.C., Sreenivas, R., & Vakili, P. (1992). Ordinal optimization of discrete event dynamic systems. *J. DEDS*, *2*(2), 61–88.

Jang, C.S.R., Sun, C.T., & Mizutani, E. (1997). Neural-fuzzy and soft computing. Englewood Cliffs, NJ: Prentice-Hall.

Lau, T.W.E., & Ho, Y.C. (1997). Universal alignment probabilities and subset selection for ordinal optimization. *JOTA*, *39*(3).

Lee, L.H., Yang, M.S., & Ho, Y.C. (1997). On stochastic optimization and its applications to manufacturing. *Lecture Notes in Applied Mathematics Series*, vol. 33 (pp 317–331).

Ruml, W., Ngo, J.T., Marks, J., & Sheiber, S. (1997). Easily search encoding for number partitioning. *JOTA*, to appear.

Papadimitriou, C.H., & Tsitsiklis, J.N. (1986). Intractable problems in control theory. *SIAM J. Control Optim. 24*(4), 639–654.

Patsis, N., Chen, C.H., & Larson, M. (1997). Parallel simulation of DEDS. *IEEE Trans. Control Technol*, to appear.

Wieseltheir, J.E., Barnhart, C.M., & Ephremides, A. (1995). Ordinal optimization of admission control in wireless multihop voice/data network via standard clock simulation. *J. DEDS*, *5*(2–3).

Witsenhausen, H.S. (1968). A counterexample in stochastic optimum control. *SIAM J. Control*, *6*(1).

Xie, X. (1997). Dynamics and convergence rate of ordinal comparison of stochastic discrete event systems. *IEEE Trans. Automat. Control*, to appear.

Mei (May) Deng received her BEng in electrical engineering from Tsinghua University (China) and her MS and PhD in engineering sciences from Harvard University. She joined AT&T (Bell) Labs in 1995 after graduation from Harvard. She has been a senior technical staff member since then. Her research interests include network design, internet applications, data network performance, simulation and optimization.

Yu-Chi (Larry) Ho received his SB and SM degrees in electrical engineering from MIT and his PhD in applied mathematics from Harvard University. Except for three years of full time industrial work he has been on the Harvard faculty. Since 1969 he has been Gordon McKay Professor of engineering and applied mathematics. Since 1989, he has been the T. Jefferson Coolidge Chair in applied mathematics and Gordon McKay Professor of systems engineering at Harvard. He was also the visiting profesor to the Cockrell Family Regent's Chair in engineering at the University of Texas, Austin in 1989.

He has published over 140 articles and three books, one of which (co-authored with A.E. Bryson, Jr.) has been translated into both Russian and Chinese and made the list of Citation Classics as one of the most referenced works on the subject of optimal control. He is on the editorial boards of several international journals and is the editor-in-chief of the international Journal of Discrete Event Dynamic Systems. He is the recipient of various fellowships and awards including the *Guggenheim (1970) and the IEEE Field Award for Control Engineering and Science (1989), the Chiang Technology Achievement Award (1993). He is a Life fellow of IEEE, a Distingushed Member of the Control Systems Society, and was elected a member of the US National Academy of Engineering (1987).* In addition to service on various governmental and industrial panels, and professional society administrative bodies, he was the President of the IEEE Robotics & Automation Society in 1988 and co-founder of Network Dynamics. Inc., a software firm specializing in industrial automation.

His research interests ranges from differential games, information structure, multi-person decision analysis, to incentive control, and since 1983, exclusively in discrete event dynamic systems, perturbation analysis, ordinal optimization, and more recently computational intelligence – topics which lie at the intersection of control system theory, operations research, and artificial intelligence.