Option Pricing with Genetic Algorithms: A Second Report *

Shu-Heng Chen
AI-ECON Research Group
Department of Economics
National Chengchi University
Taipei, Taiwan 11623
TEL: 886-02-9387308
FAX: 886-02-9390344
E-mail: chchen@cc.nccu.edu.tw

Woh-Chiang Lee
AI-ECON Research Group
Department of Business Administration
Van Nung Institute of Technology
Chung-Li, Taiwan 320
TEL: 886-03-4515811-277
FAX: 886-03-4531297
E-mail: david@cc10.vit.edu.tw

Abstract

The cross-fertilization between artificial intelligence and computational finance has resulted in some of the most active research areas in financial engineering. *One direction is the application of* machine learning techniques *to pricing financial products, which is certainly one of the most complex issues in finance. In the literature, when the interest rate, the mean rate of return and the volatility of the underlying asset follow general stochastic processes, the analytical solution is usually not available. Over the last two years, artificial neural nets have been applied to solve option pricing numerically. However, so far, there is no applications based on evolutionary computation in this area. In this paper, we shall illustrate how* genetic algorithms *(*GAs*), as an alternative to neural nets, can be potentially helpful in dealing with option pricing. In particular, we test the performance of basic genetic algorithms by using it to the determination of prices of* European call options, *whose exact solution is known from* Black-Scholes option pricing theory. *The solutions found by basic genetic algorithms are compared with the exact solution, and the performance of* **GAs** *is evaluated accordingly.*

1 Introduction

One of the most difficult issues in finance is the valuation of complex financial products, such as *financial derivatives*.[1] This valuation usually requires knowledge of the statistics of the *underlying security*, such as the *mean return* and *standard deviation of the return* (the *volatility*). If *these parameters are constant*, and if the *interest rate* is constant, then Black and Scholes (1973) has established a formula (an exact solution) for determining the value of the *European-style option*.[2] However, if these parameters are not constant and other styles of options, such as the *American-style option*, are considered, then the exact solution may not be feasible. In this case, several numerical techniques based on massive computation have been developed. Recenlty, techniques from *machine learning*, such as *artificial neural networks* (**ANNs**), have also been used to derive exact numerical solutions to option pricing (Hutchinson, et al, 1994; Barucci, et al, 1995, Lajbcygier, et al, 1996, Liu, 1996; White, 1996).[3]

This paper, to our best knowledge, is the first application of *evolutionary computation paradigm*, to option pricing. The particular style of evolutionary computation considered in this paper is *genetic algorithms*. As we learned from the recent discussion on *no-free-lunch* (**NFL**) *theorem* (Wolpert and Macready, 1995), when

*Research support from NSC grant No.85-2415-H-004-001 is gratefully acknowledged. This paper is an abbreviation of Chen and Lee (1996).

[1]A financial derivative is a financial instrument that is based on another more elementary financial instrument. The value of this financial derivative obviously depends on the price of the instrument on which the derivative is based.

[2]The solution is obtained from the equation known as the *Black-Scholes partial differential equation*.

[3]For those who are not familiar with the field *option pricing*, we shall make a remark here. The *price* referred in this field is the price defined by a *no-arbitrage partial differential equation*. Whether this price reflects the true value of the asset or whether the assumptions used to derive this partial differential equation hold in the real world is a philosophical issue known as the *efficient market hypothesis* in finance. We shall not dwell on this issue here. For recent reflections on this hypothesis, the interested reader is referred to Chen and Yeh (1996), and Chen and Tan (1996).

0-7803-4122-8/97 $10.00©1997 IEEE

comparing two machine learning tecniques, instead of asking "Which tecnhique is best?", the right question to ask should be "*When is each technique at its best?*". Motivated by the NFL theorem, this paper provides an *initial attempt* to the test whether we can benefit more by using genetic algorithms, instead of using artificial neural networks, to solve option pricing issue. As there is no unique way to apply ANNs to option pricing, the way to apply genetic algorithms to option pricing is also not unique. So, by "inital attempt", we mean that this work is just a *straightforward applications* of GAs to option pricing. At the end of this paper, some thoughts on enriching this initial attempt are provided. The rest of the papaer is organized as follows. The Black-Scholes option pricing model is briefly reviewed in Section 2. In Section 3, the relevance of genetic algorithms to option pricing, by means of the *function approximation theorem*, is shown. Section 4 discusses the parameters set in this study and describes the simulations based on GAs. The results are compared with the exact solution from the Black-Scholes partial differential equation.

2 The Black-Scholes Option Pricing Model

Black and Scholes (1973) were the first to provide a closed-form solution for the valution of *European options*. The Black-Scholes option pricing model is based on the principle known as the *no-arbitrage condition* in economics. Given a few asumptions, Black and Scholes recognized that it is possible to form a *risk-free hedge portfolio* consisting of a long position in the *stock* and a short position in the *European call* written on that stock, and that derives the following *Black-Schole partial differential equation*.

$$\frac{\partial C}{\partial t} = r_f C - r_f S \frac{\partial C}{\partial S} - \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\sigma^2 S^2 \qquad (1)$$

where S the price per share, C the price per call, t the time to buy call option, r_f the *risk-free rate*. and σ the volatility of the stock price (σ).

Equation (1) can be solved with the following two boundary conditions:

$$C(S,\tau) = MAX(S-E,0) \qquad (2)$$

and

$$C(S=0,\tau) = 0 \qquad (3)$$

Equation (2) simply says, at expiration ($t = T, \tau = 0$), a call option must have a value that is equal to zero or to the difference between the stock price and the exercise price as well as the time of maturity τ whichever is greater; otherwise there will be *arbitrage opportunities* awaiting exploitation. Equation (3) says that the call option price is worthless when $S = 0$ even if there is a long time to expiry. Black-Scholes(1973) transforms the equation into the heat exchange equation from physics to find the following solution:

$$C = SN(d_1) - Ee^{-r_f \tau}N(d_2) \qquad (4)$$

where $d_1 = \frac{ln(S/E)+r_f\tau}{\sigma\sqrt{\tau}} + \frac{1}{2}\sigma\sqrt{\tau}$, $d_2 = d_1 - \sigma\sqrt{\tau}$, and $N(d)$ is the cumulative distribution function for the standardized normal distribution. Equation (4) says that the price of an option on a stock without cash dividends depends on only five directly observable variables: (1) the stock's price (S), (2) the exercise price (E), (3) the time to maturity (τ), (4) the risk-free rate of interest (r_f), (5) the volatility of the stock price (σ).

3 Use GAs to solve OPM

Following Barucci (1995), we can approximate the calling price C by C_a.

$$C_a(S,\tau) = MAX(S-E,0) + \sum_{i=1}^{N}\psi_i(\tau)\phi_i(S) \qquad (5)$$

ψ_i and ϕ_i, $i = 1,...,N$, are known analytic functions and are called *trial functions*, $C_0(S,\tau)$ is a function chosen properly to satisfy the boundary and initial conditions. The trial functions chosen to approximate $C(S,\tau)$ in this paper are the polynomial functions for $\psi_i(\tau)$ and trigometric functions for $\phi_i(S)$.[4]

$$\psi_i(\tau) = \sum_{j=1}^{k} a_{i,j}\tau^j, \qquad (6)$$

and

$$\phi_i(S) = b_i sin(i\pi S) + c_i cos(i\pi S). \qquad (7)$$

Based on the no-arbitrage condition, i.e.,

$$L(C(S,\tau)) - rC(S,\tau) = 0, \qquad (8)$$

where

$$L(C(S,\tau)) = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - \frac{\partial C}{\partial \tau}, \qquad (9)$$

We shall define the error of our approximation R in terms of the linear operator L,

$$R = L(C_a(S,\tau)) - rC_a(S,\tau) \qquad (10)$$

By the chosen trial functions, R can be derived analytically[5], and we should use R^2 as *the fitness function* in the following GAs simulations.

[4]For other choices, readers are referred to Chen and Li (1996).
[5]For details, see Chen and Lee (1996).

4 Simulation Description and Results

Table 1: The Setting of Controlling Parameters

Number of chromosomes	2500, 5000, 7500 10000,12500
Population size	25, 50, 75, 100, 125
Number of trial functions	1, 2, 3, 4, 5
Length of string	15, 30, 45, 60, 75
Selection mechanism	roulette-wheel selection
Crossover style	two-point crossover
Crossover rate	0.6
Mutation rate	0.001
Interval of parameter	[-1,1]
Fitness function	R^2

The software used in this paper is *GENESIS 5.0*, written by John Grefenstette (Grefenstette, 1990) to promote the study of genetic algorithms for function optimization. The chosen parameters to run *Genesis 5.0* are shown in Table 1. The number of chromosomes is adapted for the expanding search space when the number of trial functions increases. Without this adapatation, Chen and Lee (1996) found that the result usually gets worse when N goes up. The test problem is the European call option with those parameters indciated in Table 2.

Table 2: The Parameters of the European Call Option

Stock price (S)	0.8, 0.9, 1.1, 1.2
Exercise pirce (E)	1
Time to maturity (τ)	1
Risk-free rate of interest (r_f)	0.1
Volatility of stock (σ^2)	0.1

There are 100 GA runs for each number of trial functions, and there are 100 generations for each GA run. The distribution of the GA-based call prices calculated by the best chromosome in Generation 100 is depicted in Figures 1-4. The vertical line appearing in each figure indicates the corresponding theoretical price based on the Black-Scholes formula. There are two statistics provided for each figure, namely, median and mean. In addition, mode can also be found by looking at the peak of each figure. There are some good signs of these simulation results. First of all, we can see that the true price is always covered by the distribution. Furthermore, the distribution has a tendency to shrink as N goes up.[6] Thirdly, in spite of the limited number of trial functions, for some cases, such as $S = 0.8$ and $S = 1.2$, the true prices can be well approximated by the mean, median or mode of the distribution. Given that the GAs applied in this study are very simple, these good signs are encouraging enough for further exploitation of this technique in future studies.

[6]For the case N=4 and 5, and the moments statistics of these distributions, see Chen and Lee (1996).

References

[1] Barucci, E., U. Cherubini and L. Landi (1995), "Neural Networks for Contingent Claim Pricing via the Galerkin Method," paper presented at the *First International Conference of the Society for Computational Economics*, IC^2 Institute, Austin, Texas, May 21-24, 1995.

[2] Black, F. and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, pp.637-654.

[3] Chen, S.-H. and W.-C Lee (1996), "Computational Intelligence in Option Pricing: Artificial Neural Nets, Genetic Algorithms and Galerkin Method," *Woking Paper Series 9612*, Department of Economics, National Chengchi University.

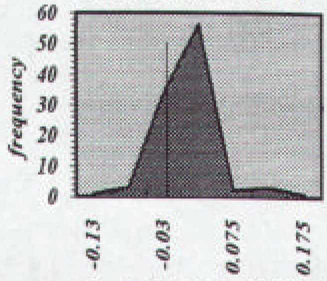
[4] Chen, S.-H. and C.-W. Tan (1996), "Measuring Randomness by Rissanen's Stochastic Complexity: Applications to the Financial Data," in D. L. Dowe, K. B. Korb and J. J. Oliver (eds.), **ISIS:** *Information, Statistics and Induction in Science*, World Scientific, Singapore. pp.200-211.

[5] Chen, S.-H. and C.-H. Yeh (1996), "Genetic Programming and the Efficient Market Hypothesis," in J. Koza, D. Goldberg, D. Fogel and R. Riolo (eds.), *Genetic Programming 1996: Proceedings of the First Annual Conference*, MIT Press, Cambridge, MA. pp. 45-53.

[6] Grefenstette,J,J(1990)," A user's Guide to GENESIS version 5.0," Mimeo, Naval REACH laboratory, Washington,D,C.

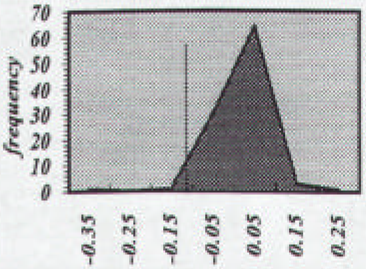
[7] Wolpert, D. H. and W. G. Macready (1995), "No Free Lunch Theorem for Search," *Technical Report SFI-TR-95-02-010*, Santa Fe Institute.

Figure 1.1:frequency of 100 experiments[S/E=0.8,N=1]



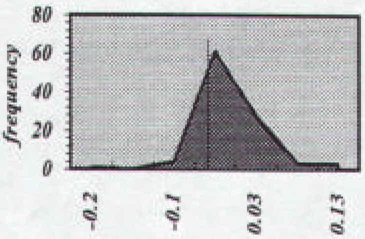
call price[B-S value=0.00536] [mean=0.0046,median=0.0006]

Figure2.1:frequency of 100 experiments[S/E=0.9,N=1]



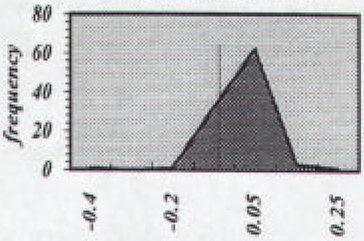
call price[B-S value=0.0221] [mean=0.0065,median=0.00039]

Figure 1.2:frequency of 100 experiments[S/E=0.8,N=2]



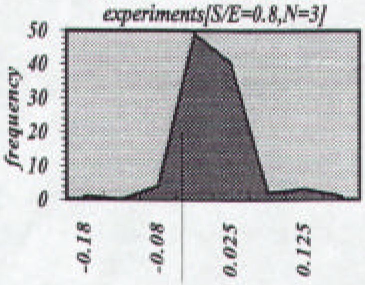
call price[B-S value=0.00536] [mean=0.0017,median=0.0015]

Figure 2.2:frequency of 100 experiments[S/E=0.9,N=2]



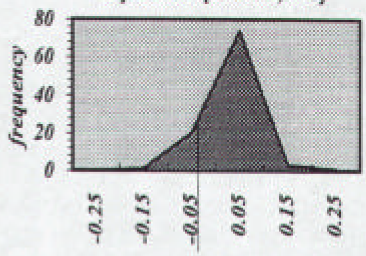
call price[B-S value=0.0221] [mean=0.0076,median=0.00089]

Figure 1.3:frequency of 100 experiments[S/E=0.8,N=3]



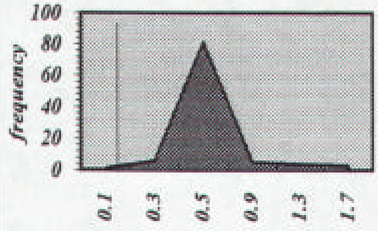
call price[B-S value=0.00536] [mean=0.009,median=-0.0004]

Figure 2.3:frequency of 100 experiments[S/E=0.9,N=3]



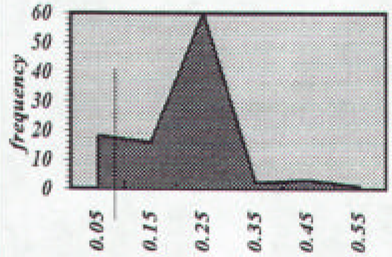
call price[B-S value=0.0221] [mean=0.0097,median=0.0034]

Figure 3.1:frequency of 100 experiments[S/E=1.1,N=1]



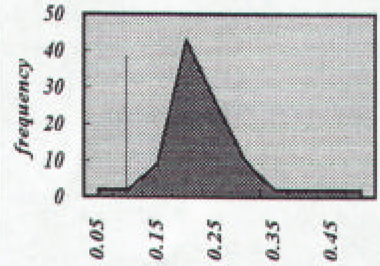
call price[B-S value=0.1121]
[mean=0.6977,median=0.0265]

Figure 3.2:frequency of 100 experiments[S/E=1.1,N=2]



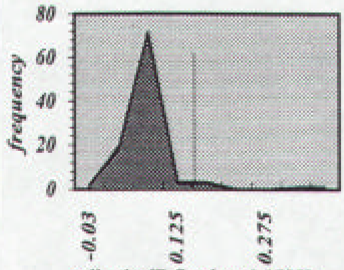
call price[B-S value=0.1121]
[mean=0.1932,median=0.2133]

Figure 3.3:frequency of 100 experiments[S/E=1.1,N=3]



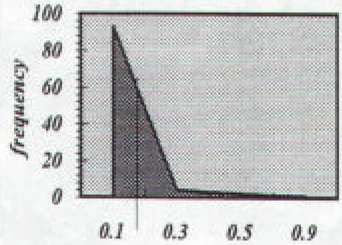
call price[B-S value=0.1121]
[mean=0.2532,median=0.2434]

Figure 4.1:frequency of 100 experiments[S/E=1.2,N=1]



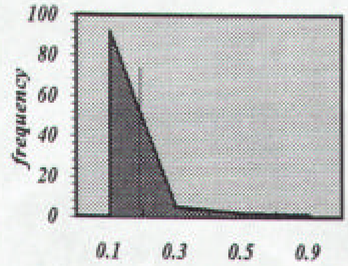
call price[B-S value=0.1757]
[mean=0.0691,median=0.0595]

Figure 4.2:frequency of 100 experiments[S/E=1.2 ,N=2]



call price[B-S value=0.1757]
[mean=0.1351,median=0.1176]

Figure 4.3:frequency of 100 experiments[S/E=1.2 ,N=3]



call price[B-S value=0.1757]
[mean=0.1553,median=0.1357]