

Evolving the Best Known Approximation to the Q Function

Dao Ngoc Phong
Dept of Information &
Communication
Hanoi City Government
Vietnam
phongaptech@yahoo.com

Constantin Siriteanu
University of Hokkaido
Japan
sircosti@gmail.com

Nguyen Xuan Hoai^{*}
IT Research &
Development Center
Hanoi University
Vietnam
nxhoai@hanu.edu.vn

Nguyen Quang Uy
Faculty of Information
Technology
Military Technical Academy
Vietnam
quanguyhn@gmail.com

R.I. (Bob) McKay
Structural Complexity
Laboratory
Seoul National University
South Korea
rimsnucse@gmail.com

Namyong Park
Structural Complexity
Laboratory
Seoul National University
South Korea
zest159@naver.com@gmail.com

ABSTRACT

The Gaussian Q-function is the integral of the tail of the Gaussian distribution; as such, it is important across a vast range of fields requiring stochastic analysis. No elementary closed form is possible, so a number of approximations have been proposed. We use a Genetic Programming (GP) system, Tree Adjoining Grammar Guided GP (TAG3P) with local search operators to evolve approximations of the Q-function in the form given by Benitez [1]. We found more accurate approximations than any previously published. This confirms the practical importance of local search in TAG3P.

Categories and Subject Descriptors

I.2.8 [Problem Solving]: Control Method—*Search, Heuristics*

Keywords

Q-function approximation, Genetic Programming, Tree Adjoining Grammar, Local Search

1. INTRODUCTION

We present Genetic Programming (GP) methods which have led to what is, under important criteria, the best approximation to the Gaussian Q-function currently known. It substantially improves in accuracy, over practically-relevant ranges, on previous approximations. It possesses desirable properties (notably, a form that is easy to integrate) that previous-best approximations do not.

^{*}Corresponding Author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO '12, July 7-11, 2012, Philadelphia, Pennsylvania, USA.
Copyright 2012 ACM 978-1-4503-1177-9/12/07 ...\$10.00.

GP [19, 13] is an evolutionary paradigm for automatically finding problem solutions. Since its inception, GP has been applied to a wide range of fields [19], and routinely exhibits human-competitive performance [14]. Tree Adjoining Grammar Guided GP (TAG3P) [6] uses tree adjoining grammars to define a language bias controlling the evolved programs (individuals). TAG3P has been demonstrated as a powerful tool in finding models approximating a given data set, especially when it is used with some readily-defined local search operators – (point) insertion and deletion [5, 7].

The Q-function is particularly important in the field of communications, where Gaussian noise is assumed. It is closely related to the cumulative density function (CDF) of the Gaussian, in which form it appears in almost every field where Gaussian distributions are used, from theoretical physics to psychology. In many scenarios, the integral is also required, because methods for computing expectations over a Gaussian typically lead to such an integral.

The Q-function's definition in the form of an improper integral makes it hard to conduct exact analyses for communication systems [2, 12]. Thus it would be highly desirable to obtain a closed (analytical) form using elementary functions. However no such solution is possible [20]. The only option has been to approximate [12]¹. A number of approximations have been proposed by mathematicians (see section 2), but the search continues.

In this paper, we discuss the use of TAG3P with local search operators to discover new approximations. To the best of our knowledge, this work (together with our previous result [17], which it substantially improves) are the first attempts to use GP to find approximations to the Q-function. We compare it with human-derived approximations.

The remainder of the paper is organised as follows. In the next section, we present the basics of Q-function approximation. In Section 3, we briefly describe tree adjoining grammars, TAG3P, and the local search operators – insertion and deletion. Section 4 explains our experimental settings. The results are presented in section 5. Section 6 briefly describes an application of our new Q-function approximation in wire-

¹The Taylor expansion gives a closed, analytical approximation, but it converges too slowly.

less communications. The paper concludes with section 7, in which we highlight some potential future extensions.

2. BACKGROUND

We define the Q-function approximation problem, and review some well-known solutions designed by human experts.

2.1 Q-function Approximation

The Gaussian Q-function is important in performance analysis for communications [21]. It is closely related to the CDF Φ of the Gaussian², and in this form its importance stretches across most applications of probability. It arises from the common assumption that system noise is Gaussian. Since we typically need to compute expectations of some property over this distribution, we need to integrate over its tail. In turn, that implies computing some form of the Gaussian error function $Q(x)$, which is defined as:

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \quad (1)$$

As Simon and Alouini [22] explain, this form is not computationally useful. What we need for computational efficiency is an approximation – ideally, in finite form (rather than a series), and preferably similar in form to the Gaussian itself (since there are known techniques for handling it).

2.2 Previous Solutions

Table 1: Historical Approximations

Name	Equation	Name	Equation
PBCS	2	OPBCS	3
CDS	4	GKAL	5
EXP	6		

A number of authors have looked for free-form approximations, with no constraints on the form. Borjesson and Sundberg [2] derived a class PBCS of approximations to the complementary error function (closely related to the Gaussian Q-function). Rewriting in terms of Q, it has the form:

$$Q(x) \approx \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}\sqrt{1+x^2}} \quad (2)$$

Chan and Beaulieu [3], call this the PBCS approximation; we follow their terminology. A similar form with optimised parameters, also proposed in [2], is:

$$Q(x) \approx \frac{1}{(1-a)x + a\sqrt{x^2+b}} \cdot \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \quad (3)$$

with $a = 0.339$, $b = 5.510$ – optimised PBCS (OPBCS).

Chiani et al. proposed a simple approximation [4] that works well for some problems, but large errors on small arguments limit its applications:

² $Q(x) = 1 - \Phi(x) = \Phi(-x)$; but most Q-function work minimises the maximum absolute error, while Φ -minimisation uses the mean relative error, due to different applications.

$$Q(x) \approx \frac{1}{12} e^{-\frac{x^2}{2}} + \frac{1}{4} e^{-\frac{2x^2}{3}} \quad (4)$$

which we denote as the CDS approximation in this paper.

The GKAL approximation of Karagiannidis and Lioumpas [12] gives smaller approximation errors. It can be written as:

$$Q(x) \approx \frac{(1 - e^{-\frac{Ax}{\sqrt{2}}})e^{-\frac{x^2}{2}}}{\sqrt{2\pi}Bx} \quad (5)$$

with $A = 1.98$ and $B = 1.135$.

Although these free-form solutions may be accurate, they may not be so easy to apply practically: what is needed is not Q itself, but a transformed form. More restricted approximations, even if less accurate, may be preferable. Benitez and Casadevall [1] proposed the EXP approximation, resembling the Gaussian, $e^{P(x)}$, with $P(x) = ax^2 + bx + c$:

$$Q(x) \approx e^{ax^2 + bx + c} \quad (6)$$

Over the range $x \in [0, 8]$, they found optimum values for a , b and c of -0.4698 , -0.5026 , and -0.8444 . It is less accurate than OPBCS, but its mathematical tractability makes it useful nevertheless [1]. The relative absolute error (RAE) of these functions is depicted in Figure 1 over the interval $[0, 8]$.

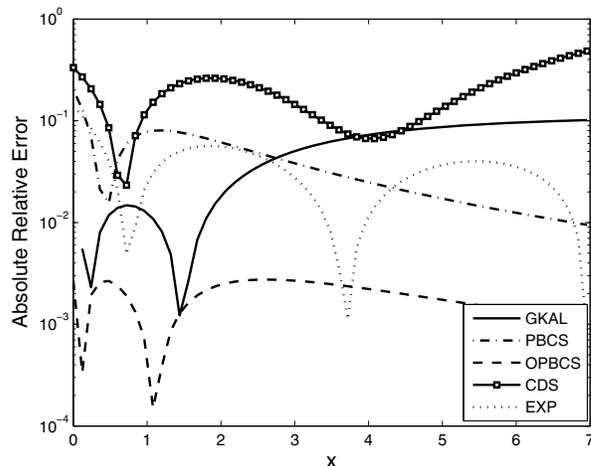


Figure 1: Relative Errors of Q-Function Approximations found by Human Experts

3. METHODS

3.1 Tree-Adjoining Grammars

Proposed by Joshi [10, 11] as tree rewrite systems, a tree-adjoining grammar (TAG) is a quintuple (Σ, N, I, A, S) :

- (i) Σ is a finite set of terminal symbols.
- (ii) N is a finite set of non-terminal symbols: $N \cap \Sigma = \emptyset$.
- (iii) S is a distinguished non-terminal symbol: $S \in N$.
- (iv) I (initial) is a finite set of finite labelled trees, known as α -trees. The labels on interior nodes are non-terminals, but on frontier nodes may be terminals. Non-terminal symbols on the frontier of an α -tree are marked as \downarrow (substitution).
- (v) A (auxiliary) is a finite set of finite labelled trees, known

as β -trees. The labels on interior nodes are non-terminals, but on frontier nodes may be terminals. One distinguished non-terminal (the foot node) on the frontier is marked with an asterisk, to denote that it is used for adjunction; all others are marked with a \downarrow (for substitution).

Adjunction builds a new (derived) tree τ' from an auxiliary tree β and another tree τ (initial, auxiliary or derived). If τ has an interior node labelled A , and β is an A -type tree (i.e. the label of its root node is A), the adjunction of β into τ to produce τ' is as follows: Firstly, the sub-tree τ_1 rooted at A is temporarily disconnected from τ . Next, β is attached to replace the sub-tree. Finally, τ_1 is attached back to the foot node of β . τ' is the final derived tree achieved from this process. Substitution also builds a new (derived) tree τ' from an auxiliary tree β and another tree τ ; but in this case, the operation is conducted on a frontier node marked with a non-terminal A . Again, β is assumed to be an A -type tree. τ' is formed by replacing the frontier node with β .

TAGs form the representation in three extant GP systems, TAG3P [5], DTAG3P [9] and TAGE [16].

3.2 Tree-Adjoining Grammar Guided Genetic Programming - TAG3P

TAG3P, proposed in [6], uses TAGs to guide GP search [6, 5]. An individual is represented by a TAG derivation tree. To evaluate fitness, the tree is decoded first into a derivation tree of the corresponding CFG, then the corresponding expression tree; fitness is evaluated on the latter as in standard GP. Figure 2 depicts the form of a TAG3P individual and its transcription to a CFG derivation tree.

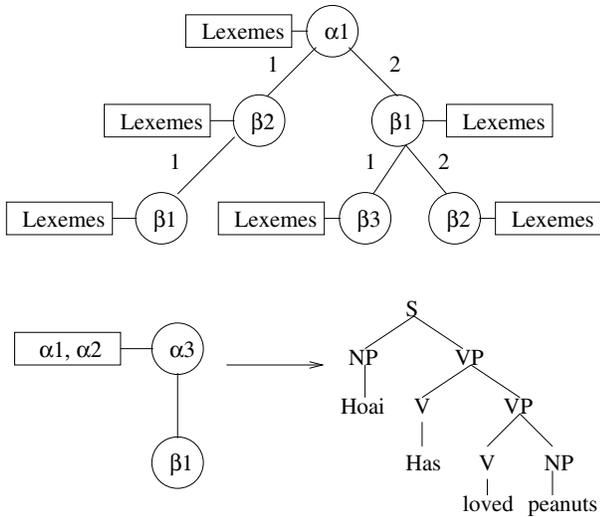


Figure 2: TAG3P Individual and its Transcription to CFG Derivation Tree

3.3 Local Search Operators

Among these operators are two local search operators, (point) insertion and deletion [5]. In insertion, a node is randomly added to the TAG-derivation tree (TAG3P individual) in an open adjoining address of a node (square node in Figures 3 and 4), while in deletion, a random node is deleted from the tree. Figures 3 and 4 show the effect of

insertion and deletion operators on a TAG3P individual.

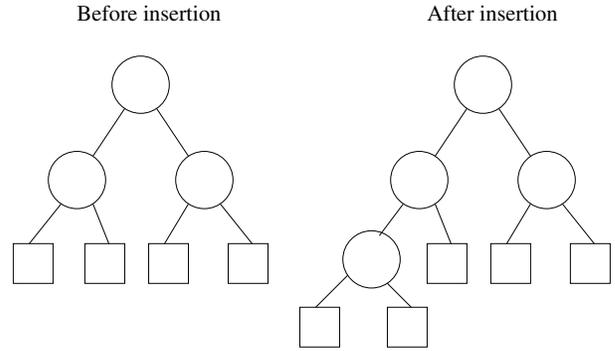


Figure 3: TAG3P Insertion Operator

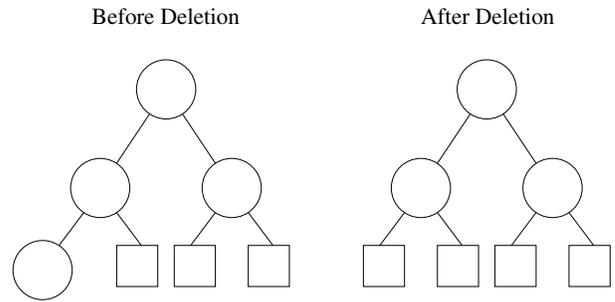


Figure 4: TAG3P Deletion Operator

In [7], it was shown that insertion and deletion are local operators (when neighbourhoods are defined by tree-edit distance). Thus these operators can act in TAG3P either as fine-grained mutation operators, or as local search operators in a hybrid evolutionary/hill-climbing search. The results of [8] strongly support the latter strategy. In this paper, we compare basic TAG3P (subtree mutation and crossover only) with a version (TAG3PL) augmented by hill-climbing local search applying the insertion and deletion operators with equal probability.

4. EXPERIMENTAL SETTINGS

Traditionally, for symbolic regression problems, the fitness function is the sum of the absolute (or sometimes the square) error. For Q-function approximation, the literature uses relative absolute error (RAE) (e.g. [2, eq. 15]), and we follow this practice. Formally, the fitness of an individual is:

$$\text{RAE} = \frac{1}{N} \sum_{i=1}^N \frac{|f_i - y_i|}{f_i} \quad (7)$$

where N is the number of data samples (fitness cases), f_i is the value of the Q-function, and y_i is the function value of the individual at the i^{th} point in the sample set. In this paper, the sample (data) set consists of 400 equidistant points over the interval $[0,8]$, together with their function values computed with Equation 1. The limited sampling interval reflects the Q-function's exponential convergence to 0 for $x > 8$: the region beyond attracts much less interest [21].

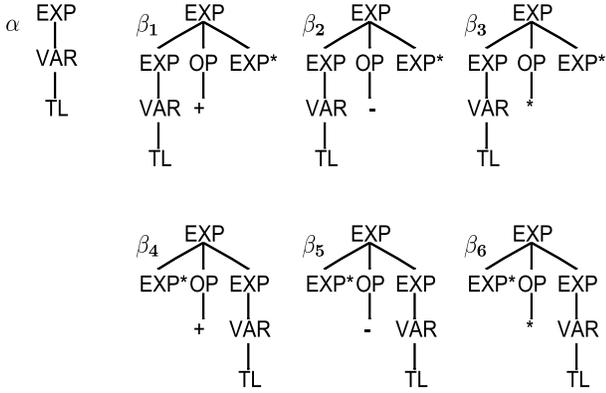


Figure 5: TAG for Q-function Approximation

Table 2: TAG3P/TAG3PL Parameter Settings

Parameter	Value
Generations	100
Population size	3000 (TAG3P) 100 (TAG3PL)
Selection	Tournament
Tournament size	9
Crossover probability	0.9
Local search operator	Randomly selected Insertion or Deletion Operator
Local search steps	30 (TAG3PL) 0 (TAG3P)
Local search strategy	Hill-Climbing
Raw fitness	Relative absolute error over all fitness cases
Termination	Max generations exceeded
Trials per treatment	50 independent runs

Other factors that affect the performance of GP/TAG3P are the set of non-terminals and terminals. These sets generally depend on the problem. In this paper, we focus on approximations of the Q-function of the form $e^{P(x)}$, as in [1]. Thus the non-terminal and terminal sets are respectively: +, -, *, and x , ERC, π , 1^3 . Here, ERC is an ephemeral random constant that takes a value from the interval (0,1). The TAG for this problem is depicted in Figure 5.⁴ To compare TAG3P and TAG3PL, we ran the two settings with the same number of fitness evaluations. Other settings were determined by preliminary experiments, see Table 2.

5. RESULTS AND DISCUSSION

To determine potentially good solutions for the Q-function approximation problem, at the end of each run, we recorded the best-of-run individual if it had an acceptable RAE value relative to the human designed approximations. Table 3 shows the RAEs for the previous approximations (section 2), on our training data.

Based on this, we recorded two classes of solutions: those

³Since we assume here that the form of the target function is $e^{P(x)}$, the evolved part is $P(x)$

⁴TL in the Figure stands for a lexicon that can be substituted with x , ERC, π or 1.

Table 3: Previous Results

Name	RAE	Name	RAE
PBCS	0.0346417	OPBCS	0.0017471
CDS	0.2437469	GKAL	0.0614184
EXP	0.0348177		

having an RAE $\leq 10^{-2}$ (better than any human approximation except for OPBCS) and those having an RAE $\leq 10^{-3}$ (better than the best human approximation).

For TAG3P, out of 50 runs we have only one solution satisfying the first performance criterion, and none satisfying the second. That is, TAG3P is barely able to achieve human-comparable performance. In comparison, TAG3PL found 30 solutions satisfying the first criterion (i.e. human-comparable performance), and four satisfying the second (i.e. better-than-human performance). We attribute this to the ability of TAG3PL to fine-tune GP structure and parameters. Of these solutions, the best has an RAE of 0.0006189 (substantially better than OPBCS). It has the form

$$\text{Best_Exp_Form} = e^{ax^6+bx^5+cx^4+dx^3+ex^2+fx+g} \quad (8)$$

using the parameter values shown in table 4.

Table 4: Coefficients for Equation 8

Coeff.	Value	Coeff.	Value
a	0.0000018643	b	-0.000109
c	0.002238	d	-0.023735
e	-0.344644	f	-0.774128
g	-0.698740		

Following [3], we evaluate this approximation by plotting its relative error. For the sake of comparison, the relative error of the best approximation produced by human experts in the literature, OPBCS, and the approximation of similar form, EXP, are also plotted in Figure 6.

It can be seen from this figure that the accuracy of the approximation found by TAG3PL is competitive with OPBCS in terms of error in the interval (0,1.5). For $x > 1.5$ it is uniformly better than OPBCS. This is important because, from the time of its introduction in 1979, no better approximation has been found by human experts. But where this new approximation wins hands down is in tractability. Despite its worse accuracy, the EXP approximation [1] has been heavily used because of its tractability. The new form is as tractable as the EXP approximation, but at least as accurate as OPBCS – and in most ranges of interest, more accurate. Compared to EXP, it is – very substantially – more accurate almost everywhere. Thus TAG3PL has been effective in deriving highly accurate, and closed form, approximations of the Q-function. In this, it has substantially improved on the previous best, human-derived, approximations.

6. APPLICATION IN COMMUNICATIONS

In this section, we test our new approximation to the Q-function in evaluating the symbol detection performance of

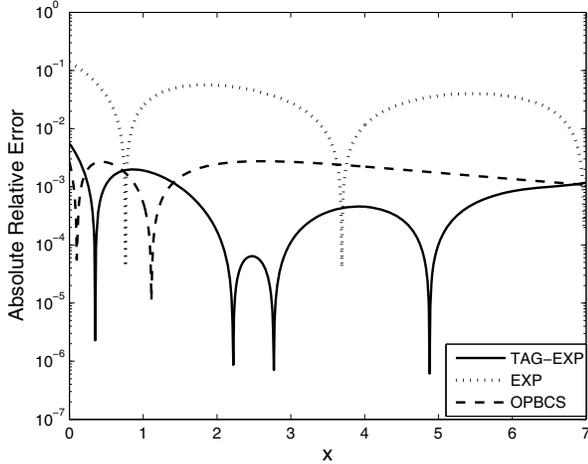


Figure 6: Relative Errors of the Approximations found by TAG3PL, EXP, and OPBCS

a wireless communications receiver. From digital communications theory it is known that the symbol error probability for many popular modulation techniques can be expressed in terms of $Q(\sqrt{\gamma})$, where γ is the signal-to-noise ratio (SNR) per symbol at the receiver [22, Chapter 8]. In order to test the accuracy of our Q-function approximation, we have selected the case of differential quadrature phase-shift keying (DQPSK) [22, Chapter 4] whose error probability has the expression [22, Eqn. (8.38), p. 203]:

$$P_e(\gamma) = 4Q(\sqrt{\gamma}) - 8Q^2(\sqrt{\gamma}) + 8Q^3(\sqrt{\gamma}) - 4Q^4(\sqrt{\gamma}) \quad (9)$$

Note that, compared to other digital modulation techniques, DQPSK has a complicated P_e expression in the Q-function. Nevertheless, our approximation can yield accurate P_e results, as shown next.

Fig. 7 shows the DQPSK symbol error probability vs. SNR (in decibels) computed with (9) by employing integration (which yields the true P_e) as well as by employing the best previously-available approximation (OPBCS) and our new TAG-based approximation. Note that only P_e values between 10^{-1} and 10^{-6} are practically relevant. This figure suggests that both approximations yield P_e near the true one. Next, we look more closely at approximation accuracy.

Fig. 8 shows the relative error in P_e approximation. Note that our new approximation is much more accurate than the OPBCS variant over the relevant SNR range. Other results (not shown here due to space limitations) obtained for other digital modulation techniques have yielded similar results. Thus, we conclude that our new approximation can help evaluate P_e much more accurately than other available approximations. Furthermore, the exponential form of our approximation helps to straightforwardly evaluate the P_e expressions of most digital modulation techniques.

7. CONCLUSION

The primary objective of this work was to improve on the solutions of a practically-important problem (Q-function approximation), which has current applications in analysis and optimisation of communications systems, and in a number of other fields. We have shown that a GP system (TAG3P)

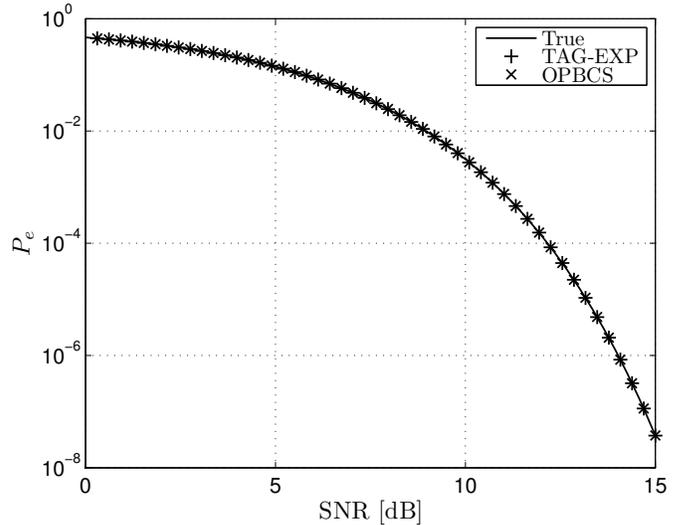


Figure 7: DPSK P_e relative error.

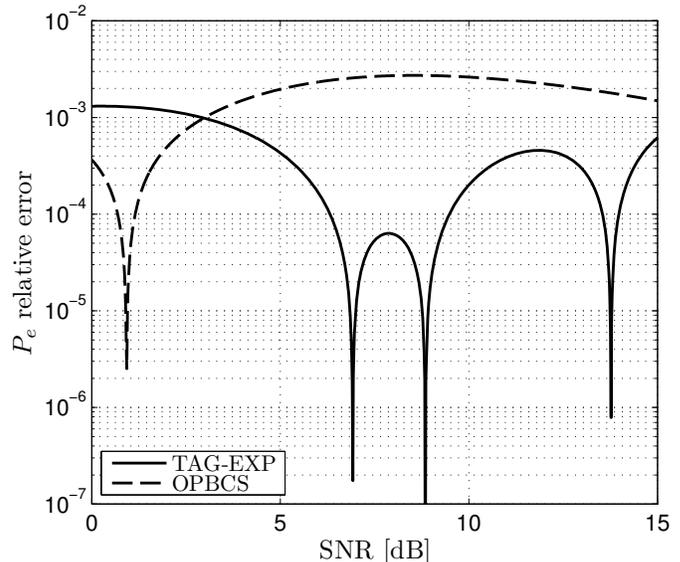


Figure 8: DPSK symbol error probability computed with (9) various formulas.

could discover new approximations that are better, in terms both of accuracy and of analytical tractability, than the best found by human experts. The best closed form approximation obtained by TAG3P is not only more accurate than the best human-derived approximation, but also easier to use in computing expectations and other computations important in practical applications. The results also confirmed the value of local search operators in TAG3P.

In future, we are planning to extend the work in this paper in a number of ways. Firstly, we hope to find more accurate approximations by giving TAG3PL more computational time. Second, we will investigate the use of bloat control methods such as Tarpeian Bloat Control [18, 15], to test whether TAG3PL can find simpler solutions of similar quality. Finally, we plan to apply these approximations in communications analysis, to build more accurate simulations of specific systems.

Acknowledgment

The work in this paper was funded by The Vietnam National Foundation for Science and Technology Development (NAFOSTED), under grant number 102.01-2011.08. The third and sixth authors were supported by the Basic Science Research Program of the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science and Technology (Project No. 2011-0004338); the Institute for Computer Technology at Seoul National University provided research facilities for this study.

8. REFERENCES

- [1] M. Benitez and F. Casadevall. Versatile, accurate, and analytically tractable approximation for the gaussian q-function. *IEEE Transactions on Communications*, 59(4):917–922, 2011.
- [2] P. Borjesson and C. Sundberg. Simple approximations of the error function $q(x)$ for communications applications. *IEEE Transactions on Communications*, 27:639–643, 1979.
- [3] Y. Chan and N. Beaulieu. A simple polynomial approximation to the gaussian q-function and its application. *IEEE Communications Letters*, 12:124–126, 2009.
- [4] M. Chiani, D. Dardari, and M. K. Simon. New exponential bounds and approximations for the computation of error probability in fading channels. *IEEE Transactions on Wireless Communications*, 2(4):840–845, 2003.
- [5] N. X. Hoai. *A Flexible Representation for Genetic Programming from Natural Language Processing*. PhD thesis, Australian Defence force Academy, University of New South Wales, Australia, 2004.
- [6] N. X. Hoai, R. I. McKay, and H. A. Abbass. Tree adjoining grammars, language bias, and genetic programming. In *Genetic Programming, Proceedings of EuroGP’2003*, volume 2610 of *LNCS*, pages 335–344. Springer-Verlag, 2003.
- [7] N. X. Hoai, R. I. B. McKay, and D. Essam. Representation and structural difficulty in genetic programming. *IEEE Transactions on Evolutionary Computation*, 10(2):157–166, 2006.
- [8] T. H. Hoang, N. X. Hoai, R. I. B. McKay, and D. Essam. The importance of local search: A grammar based approach to environmental time series modelling. In T. Yu, R. L. Riolo, and B. Worzel, editors, *Genetic Programming Theory and Practice III*, volume 9 of *Genetic Programming*, chapter 11. Springer-Verlag, 2005.
- [9] T. H. Hoang, R. B. McKay, D. Essam, and N. X. Hoai. On synergistic interactions between evolution, development and layered learning. *IEEE Transactions on Evolutionary Computation*, 15(3):287–312, June 2011.
- [10] A. Joshi, L. Levy, and M. Takahashi. Tree adjunct grammars. *Journal of Computer and System Sciences*, 10(1):136–163, 1975.
- [11] A. Joshi and Y. Schabes. Tree-adjoining grammars. *Handbook of formal languages*, 3:69–124, 1997.
- [12] A. Karagiannidis and A. Lioumpas. An improved approximation for the gaussian q-function. *IEEE Communication Letters*, 11:644–646, 2007.
- [13] J. Koza. *Genetic Programming: On the Programming of Computers by Natural Selection*. MIT Press, MA, 1992.
- [14] J. R. Koza. Human-competitive results produced by genetic programming. *Genetic Programming and Evolvable Machines*, 11(3/4):251–284, 2010. Tenth Anniversary Issue: Progress in Genetic Programming and Evolvable Machines.
- [15] S. Mahler, D. Robilliard, and C. Fonlupt. Tarpeian bloat control and generalization accuracy. In *Genetic Programming, 8th European Conference*, volume 3447 of *LNCS*. Springer-Verlag, 2005.
- [16] E. Murphy, M. O’Neill, and A. Brabazon. Examining mutation landscapes in grammar based genetic programming. In *Proceedings of the 14th European Conference on Genetic Programming, EuroGP 2011*, volume 6621 of *LNCS*, pages 130–141. Springer Verlag, 2011.
- [17] D. N. Phong, N. Q. Uy, N. X. Hoai, and R. I. B. McKay. Evolving approximations for the gaussian q-function by genetic programming with semantic based crossover. In *Congress on Evolutionary Computation*, page To Appear. IEEE, IEEE Press, June 2012.
- [18] R. Poli. A simple but theoretically-motivated method to control bloat in genetic programming. In *Genetic Programming, 6th European Conference*, volume 2610 of *LNCS*, pages 204–217. Springer-Verlag, 2003.
- [19] R. Poli, W. Langdon, and N. McPhee. *A Field Guide to Genetic Programming*. <http://lulu.com>, 2008.
- [20] M. Rosenlicht. Liouville’s theorem on functions with elementary integrals. *Pac. J Math.*, 24(1):153–161, 1968.
- [21] M. Simon. *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*. Kluwer Academics, 2002.
- [22] M. Simon and M. Alouini. *Digital Communications Over Fading Channels: A Unified Approach to Performance Analysis*. Wiley and Sons, 2000.