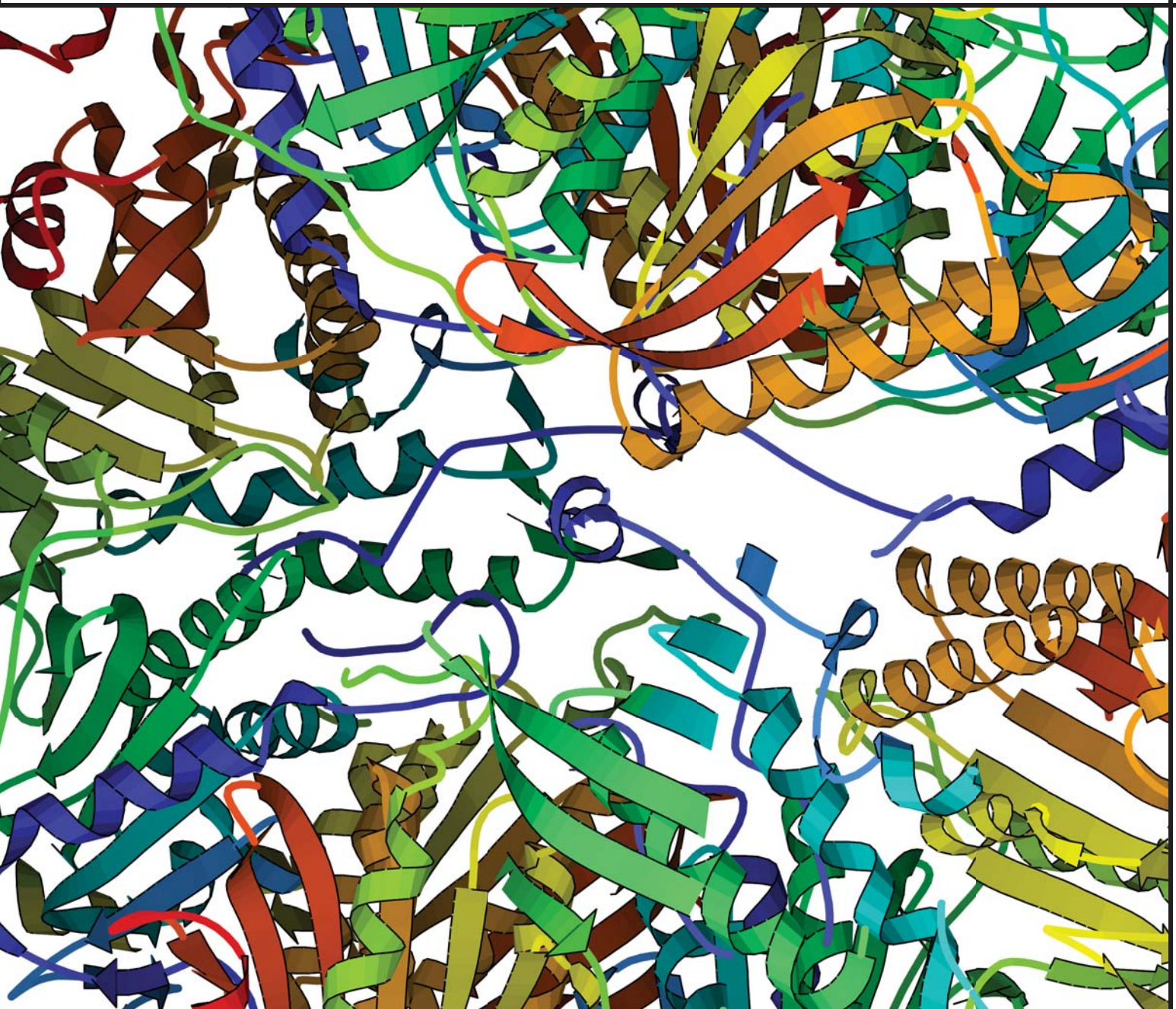
SIGEVOlution

newsletter of the ACM Special Interest Group on Genetic and Evolutionary Computation

Spring 2008
Volume 3 Issue 1



in this issue

Editorial

I always feel energized after GECCO. There are so many interesting tutorials and so many interesting presentations that, as soon as I am back home, I feel the urge to start working right away. GECCO is truly an invaluable source of inspiration. This perfect tuned engine is the fruit of the dedication and the commitment of many people who work a year around to give us this basically perfect mixture of research, social-networking and entertainment. In Atlanta, I had my first glimpse of what happens behind the scene every year and I have been astonished by the amount of work the organizers deal with to bring us GECCO. I also realized that, while the committee changes every year, Pat Cattolico is always there to help the newly appointed organizers keeping up with the many tasks, the several deadlines, and the endless details. GECCO owes her a lot and I am sure everybody who attended the conference even just once will agree with me. Thank you Pat!

This is the first issue of the third volume and, believe it or not, the second issue is already on its way. We have almost caught up with the delay we accumulated during 2007. As always, we did our best to pack it with interesting content and we hope you will enjoy reading it. In the first paper, Jaume Bacardit, Michael Stout, and Natalio Krasnogor tell us how estimation of distribution algorithms can be applied to simplify protein representation. In the second paper, Cem Baydar shows how agent-based simulation and evolutionary computation can team up to develop personalized pricing policies which can beat the one-size-fits-all loyalty programs. Then, in a letter, William Langdon comments on the future of academic publishing based on his recent experience with the new book he coauthored, "A Field Guide to Genetic Programming", which has been made available on-line free of charge as a PDF to download. The usual columns complete the issue providing information about new software, the CIGPU workshop at WCCI-2008, and the forthcoming events.

The cover image was produced using the KiNG visualizer, a tool created in the Richardson lab at Duke University, and a protein taken from the RCSB Protein Data Bank.

As always, I owe thanks to the people who made this possible, Jaume Bacardit, Mike Stout, Natalio Krasnogor, Cem Baydar, William B. Langdon, Douglas A. Augusto, Patrick O. Stalph, Martin V. Butz, Garnett Wilson, Simon Harding, Francesco Amigoni, Mario Verdicchio, Ester Bernadó, Cristiana Bolchini, Ying-Ping Chen, Tian-Li Yu, Marc Schoenauer, Stewart Wilson, and board members Dave Davis and Martin Pelikan.

Pier Luca
August 14th, 2008

SIGEVOlution
Spring 2008, Volume 3, Issue 1

Newsletter of the ACM Special Interest Group on Genetic and Evolutionary Computation.

Contributors to this Issue

Jaume Bacardit
Michael Stout
Natalio Krasnogor
Cem Baydar

Contents

ISSN: 1931-8499

A Tale of Human-Competitiveness in Bioinformatics

Jaume Bacardit, Michael Stout, & Natalio Krasnogor
ASAP research group, School of Computer Science, University of Nottingham, UK

A key open problem, which has defied scientists for decades is the problem of predicting the 3D structure of proteins (Protein Structure Prediction - PSP) based on its primary sequence: the amino acids that compose a protein chain. Full atomistic molecular dynamics simulations are, for all intents and purposes, impractical as current empirical models may require massive computational resources. One of the possible ways of alleviating this cost and making the problem easier is to simplify the protein representation based on which the native 3D state is searched for. We have proposed a protocol based on evolutionary algorithms to perform this simplification of the protein representation. Our protocol does not use any domain knowledge. Instead it uses a well known information theory metric, Mutual Information, to generate a reduced representation that is able to maintain the crucial information needed for PSP. The evaluation process of our method has shown that it generates alphabets that have competent performance against the original, non-simplified, representation. Moreover, these reduced alphabets obtain better-than-human performance when compared to some classic reduced alphabets.

Introduction

Proteins are crucial molecules for the proper functioning of living organisms. Understanding how they work can help humanity explain some of the still unsolved mysteries of life. The first step towards this solution is knowing the complex 3D structure of proteins. Proteins are composed by chaining together some molecules known as amino acids (AA)[1]. The AA composition of a protein (known as primary sequence) is relatively easy to know. However, this chain folds to create the complex 3D structure of a protein, which is difficult to determine experimentally. Therefore, this structure has to be predicted from the AA composition of the proteins, and this is called the protein structure prediction (PSP) problem. Despite many decades of research in PSP, this problem remains unsolved. Quite good techniques exist for different subsets of proteins, but there is no overall good solution. Moreover, PSP is computationally a very costly process. One of the currently best PSP methods, Rosetta@home [18], used a massive collaborative computing system to predict protein structures, dedicating up to 10000 computing days to predict the structure of a single protein.

One of the ways in which this computational cost can be reduced is by simplifying the representation of the proteins that has to be explored to obtain the models for their 3D structure. There are 20 possible AA types that can appear in proteins, thus, we can define a protein chain as being a string drawn from a 20-letter alphabet. These AA types can be characterized by various physico-chemical properties, and different groups of AA types share some of these properties. Thus, it would make sense in order to simplify the protein representation to create a new alphabet where the AA types that share some properties are all identified by the same letter. In this way we would reduce the total number of letters of the alphabet and hence the complexity of the problem being solved. This process is known as *alphabet reduction*, and can benefit the prediction of several PSP subproblems that are important milestones towards a full 3D prediction of a protein structure. As in any simplification process, alphabet reduction has to be done very carefully in order to avoid losing crucial information required to predict properly the 3D structure of proteins.

[1] Also frequently called *residues*

In [4] we proposed an automated alphabet reduction protocol based on evolutionary algorithms that can be applied to simplify the representation of a variety of PSP subproblems. This protocol *tailors* the alphabet reduction specifically to the subproblem that has to be solved because, as we will show later, different PSP subproblems need different reductions. This protocol was not based on any domain knowledge to perform the alphabet reduction process. Instead, we used a well-known information theory metric, Mutual Information (MI) [8], to identify the reduced alphabet that manages to maintain as much as possible the crucial information needed to predict the PSP subproblem being solved. As an initial proof of concept we have applied this protocol to one PSP subproblem, comparing our method against the original AA alphabet. Our method is able to generate alphabets of reduced sizes that obtain similar performance to the original AA alphabet, and obtains better performance than some classic human-proposed reduced alphabets. Thus, this automated alphabet reduction protocol is human-competitive, it is applied to a very relevant problem, and it manages to achieve its objective (of reducing the complexity of the problem) without significant information loss.

Proteins and Protein Structure Prediction

Proteins are essential molecules for the functioning of life, having a variety of functions. They can take part of the structure of organisms (e.g. skin, hair), catalyze process (enzymes), transport substances (haemoglobin), or take part in the immune system of species (e.g. the immunoglobulin family of proteins), among other functions. The human genome project has provided millions of protein sequences. However, we only know the 3D structure of a small fraction of them. The sequence for millions of non-human proteins is known too. Having accurate knowledge of the 3D structure of proteins is crucial as this structure determines the function that each protein has. By understanding the exact function of proteins (and how this function is carried out) we can have a better understanding of the general mechanisms of life. Hence, the need to predict the 3D structure of proteins from their primary sequence. Another consequence of having better models of proteins is the ability to engineer proteins with higher chances of working properly. This can lead to better genetic therapy methods, synthesis of drugs for incurable diseases, improved crop production, etc. Thus, PSP is a very relevant problem with high impact on society. For instance, it was identified as a Grand Challenge by the USA government [1].

We do not know exactly how proteins fold, but it is thought that this folding process has several steps. The first step, called secondary structure, consists of some "patterns" created due to local interactions of the AAs with their nearest neighbours in the chain. Some of these patterns are called alpha helix and beta sheets. These local structures can group in several conformations or domains forming a tertiary structure. Secondary and tertiary structure may form concomitantly. The final 3D structure of a protein consists of one or more domains. Figure 1 illustrates this process.

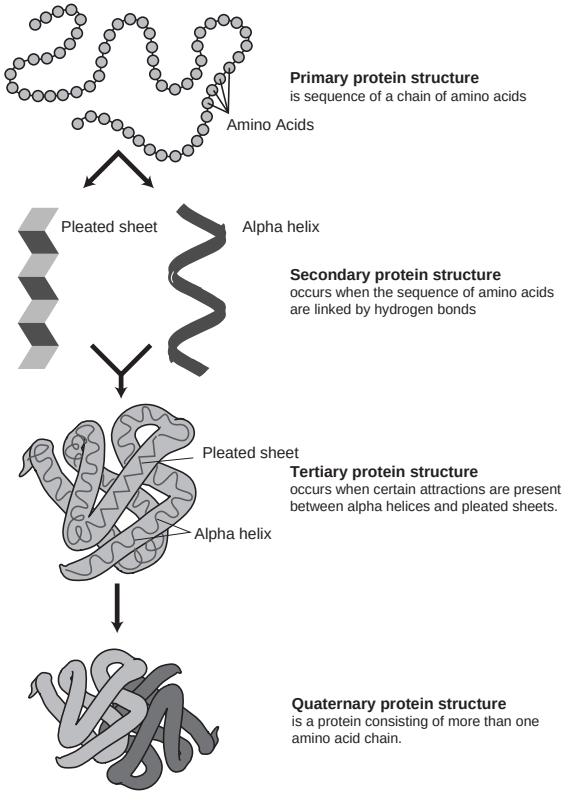


Fig. 1: Graphical representation of protein folding. Top: residues in the unfolded chain are represented by a chain of circles. Next, residues begin to form contacts. Short range contacts lead to formation of helical and pleated sheet structures. Finally the overall folded structure is formed. (Illustration Courtesy of National Human Genome Research Institute)

PSP can be tackled in many different ways. One of the possible ways is by using a divide-and-conquer approach where the problem of predicting the tertiary structure of a given sequence is split up into smaller challenges of predicting separately some structural features for a protein chain. The predictions of these features are combined afterwards to constrain the conformation space that has to be explored in the overall PSP process. Some of these features are, for instance, the secondary structure pattern that each amino acid in a protein takes, or the prediction of the ratio of surface of an amino acid that is exposed to the environment of the protein, known as solvent accessibility (SA). A third feature is called contact number (CN). In the native state each residue will have a set of spatial nearest neighbours. The number of nearest neighbours of a given residue is its contact number. This metric is a simplified profile of the end product of the protein folding process. Other structural features and alternative topology-based definitions of contact have also been investigated [23, 25, 3].

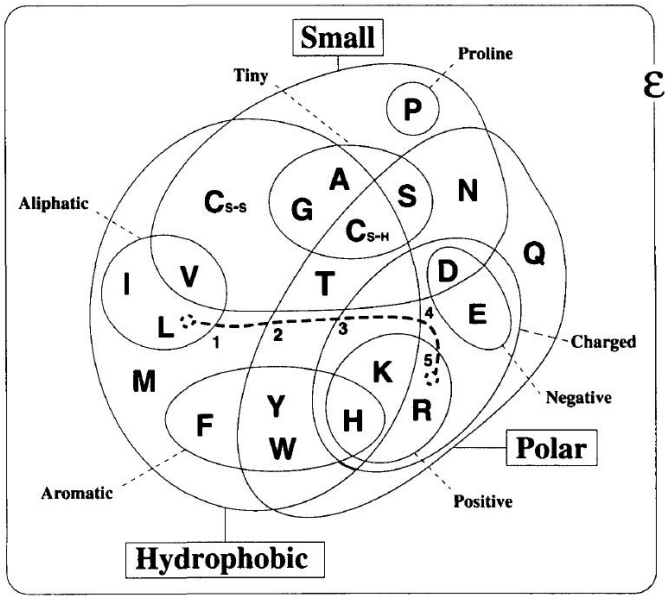


Fig. 2: Amino Acid Properties. A simplified overview of the physical and chemical properties of the amino acids. From Livingstone, C.D., Barton, G.J.: Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. Computer Applications in the Biosciences 9 (1993) 745-756, by permission of Oxford University Press

Alphabet Reduction in PSP

As we have said in the introduction section, one of the ways in which we can simplify the protein representation used for solving several PSP subproblems is to cluster the 20 amino-acid types into a small set of groups. The benefits for this process are a faster and potentially easier exploration process. Clustering together amino acid types makes sense, a priori, because amino acids have various physico-chemical properties, and some of them are shared between amino acids. Figure 2 shows a Venn diagram of some of these AA properties. We have used these properties to visualize the results of our automated protocol later in the paper.

An example of a widely explored alphabet reduction option is to transform the 20 letters AA alphabet into a two letters hydrophobic/polar (HP) alphabet. This reduction is usually followed by constraining the residue locations of the predicted protein to those of a 2D/3D lattice [13, 28, 12, 9, 14]. Figure 3 shows a simplified version of the 3D structure of a protein, where each amino acid is represented by a sphere. The protein is represented twice, in one of them each AA type has a different color. In the other one all hydrophobic residues have red color, while all polar residues are blue. Different scales of assigning AA types to either hydrophobic or polar state exist [6, 16] as well as real-valued hydrophobicity scales [7]. Some of these scales were human-designed, and as we will show later in this paper, our protocol is able to automatically generate alternative scales (without any human intervantion nor domain knowledge), tailored specifically for the problem at hand, giving higher performance.

The HP alphabet, while widely explored, is usually a too simple representation. Too much crucial information is lost in the simplification process. Thus, more recent works in alphabet reduction for PSP aim at finding alphabets of four or five letters [27, 21, 19, 17].

Automated Alphabet Reduction with Evolutionary Algorithms

For the last three years we have been applying Genetics-Based Machine Learning (GBML) techniques to solve a variety of PSP subproblems [22, 24, 5, 4, 25, 23], such as the mentioned CN and SA, and we have even proposed a new structural feature, the Recursive Convex Hull (RCH) [23], that is able to capture complementary information to CN and SA (among other PSP subproblems).

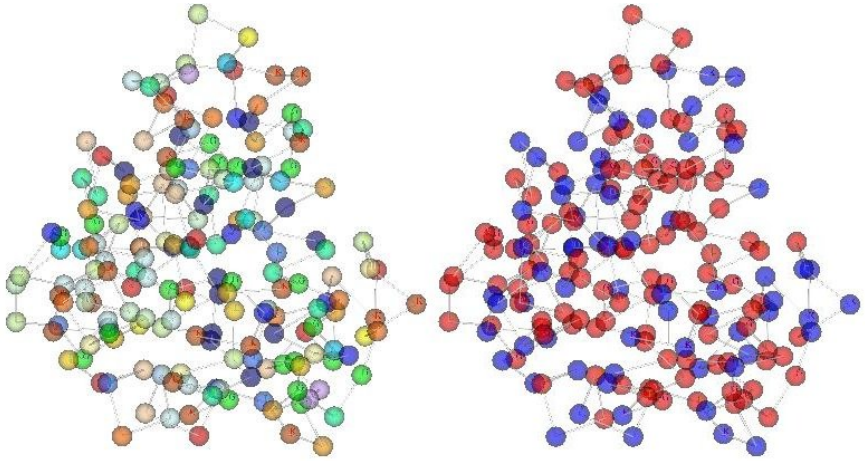


Fig. 3: Simplified visualization of a protein using either the 20-letter AA alphabet or the two-letter HP alphabet.

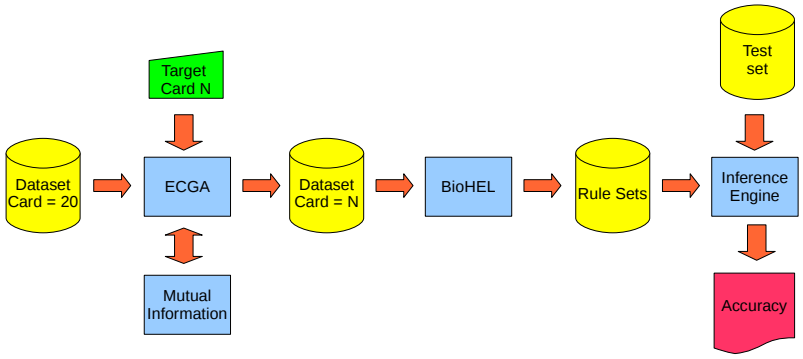


Fig. 4: Workflow of our automated alphabet reduction protocol

We have constructed an online server placed at www.infobiotic.net where rule sets generated by our GBML methods can be queried to predict many different structural features of proteins.

Initially we used GAssist [2] a Learning Classifier System [11, 20] using the Pittsburgh approach. This system generates accurate, compact and highly interpretable sets of rules. GAssist was able to obtain good results on some of these PSP datasets. However, we observed that its scalability was limited on the larger datasets. In order to overcome these limitations we created a new system, called BioHEL (Bioinformatics-oriented Hierarchical Evolutionary Learning). This system follows the Iterative Rule Learning approach first used in GBML by Venturini [26]. BioHEL contains several efficiency enhancement techniques that combined with various kinds of ensemble learning techniques allow it to successfully solve large-scale datasets such as some of the PSP sub-problems, with hundreds of thousands of instances and up to hundreds of attributes. Our GBML methods give accurate and competent solutions, but their computational cost is still quite high. Thus, we have chosen to apply alphabet reduction to our protein representation to alleviate this cost.

Automated alphabet reduction protocol

Our automated alphabet reduction protocol follows the workflow displayed in Figure 4. The initial data is the dataset predicting some PSP sub-feature having cardinality 20 (the AA types) and a target size N for the reduced alphabet. All this data is fed into the optimization algorithm that is going to find the best reduced alphabet. As optimization algorithm we have used the Extended Compact Genetic Algorithm (ECGA) [10], a method belonging to the Estimation of Distribution Algorithms (EDA) [15] family. ECGA needs a fitness function to guide its process of finding the best reduced alphabet. The goal of such functions is to identify the reduced alphabet that is able to keep all the crucial information (or as much as possible) necessary to predict our target PSP sub-feature. Ideally, we could simply use a learning algorithm applied to the dataset with reduced alphabet, but this would be very costly. Thus, we need a *cheaper* estimator of information content. We have chosen a well known information theory metric, the Mutual Information (MI) [8], for such task. MI is a measure of the interrelationship between two variables. In this case these two variables are (1) the input data (with reduced representation) used to predict our feature and (2) the feature itself. Informally we could say that the task of MI is to quantify how much the reduced input information can tell about the target feature.

ECGA produces as a result of its optimization process a transformed dataset using the optimal reduced alphabet. The next step is to verify if this process has been correct (the reduced alphabet is able to capture the necessary domain information) or not. To do so, we will learn the PSP subproblem using the reduced alphabet and compare the obtained accuracy against the accuracy obtained from the original 20-letter alphabet. To do so we have employed BioHEL. After the learning process, BioHEL generates an ensemble of rule sets. This ensemble will be fed with the test sets of a 10-fold cross-validation process to provide an accuracy measure. The comparison of this accuracy against that obtained from the 20-letter alphabet will tell if the alphabet reduction process has been successful or not.

Results

As an initial proof-of-concept of this protocol we predicted one PSP subfeature, namely CN, using a dataset of 1050 proteins and almost 260000 residues, optimizing alphabets of two, three, four and five letters. We did not try to generate alphabets of larger size because other works in the literature also focus on alphabets of similar sizes. We compared the accuracy obtained by our reduced alphabets against the accuracy obtained from the original 20-letter alphabet. Table 1 contains the results of this comparison. We also compared the accuracy of the solutions (rule sets) obtained from each alphabet using two metrics: number of rules and number of expressed attributes in each rule. The accuracy results reported in the paper are computed using the protein-wise accuracy metric. In this metric, the prediction accuracy is computed separately for the amino acids in each protein chain, and then averaged across chains. In this way, the obtained accuracy is not biased towards longer chains.

First of all, we can see how the solutions generated when learning from the reduced alphabets are always much more compact and simple as reflected by the complexity metrics. In relation to the accuracy obtained by each reduced alphabet, we can extract different observations. First of all, the most reduced alphabet (of size 2) obtains an accuracy which is 1.2% lower than the accuracy from the original alphabet. In previous work [24] we compared the performance of the AA alphabet against the most popular human-designed two-letter alphabet, the Hydrophobic-Polar alphabet [6] also for CN prediction. In those experiments, the performance gap between the HP alphabet and the full AA alphabet was 3.8%.

Thus, with our automated protocol we have been able to reduce more than three times the performance gap between the simplest possible alphabet and the original representation.

Although the automatically generated two-letter alphabet obtains better results than our previous work, its performance its still significantly worse than the performance of the AA alphabet, according to Student t-tests with 95% confidence level. Thus this reduction, as we expected, is too large and critical information (to predict CN) has been lost in the process. It would be expected that larger alphabet sizes were able to reduce the performance gap. Indeed this is what happens, if we look at the results of the three-letter alphabet. This alphabet managed to reduce the performance gap to a, non significant, difference of 0.6%. However, alphabets of sizes larger than three letters had their performance degraded again, specially in the case of the five-letter alphabet. The reason for this issue is a well known problem of the mutual information metric when applied to datasets of small sample size, degrading the robustness of the metric. This problem is explained in depth in [4].

What is the composition of the reduced alphabets generated by our protocol? Table 2 shows the alphabets of two and three letters optimized for contact number prediction. We have decided not to show the four and five-letter alphabets because they are mainly artifacts, due to the problem of the mutual information metric that we have mentioned above. We have colored each amino acid type according to various physico-chemical properties. We can observe that the two-letter alphabet it is indeed an HP alphabet, separating hydrophobic from polar residues. However, this alphabet has been automatically *tailored* to keep the crucial information for the problem at hand (CN prediction).

#letters	PWA	#rules	#expr. att./rule
Orig.	77.0±0.7	22.5±1.8	8.88±0.34
2	75.8±0.7•	11.3±0.6	5.39±0.49
3	76.4±0.7	16.7±1.4	5.95±0.98
4	76.1±0.8	15.4±1.3	6.18±1.17
5	75.7±0.8•	14.6±1.5	6.93±1.05

Tab. 1: Protein-wise accuracy, average rule set size and average number of expressed attributes per rule of BioHEL applied to the reduced datasets. • marks the cases where the reduced dataset had significantly worse performance than the original dataset with AA type representation.

#letters	Groups of letters
2	CLVIMAFYWGH/TSNRKDEPQX
3	CLVIMAFYW/GHTS/NRKDEPQX
FWY - aromatic, neutral, hydrophobic; ACILMV - hydrophobic; DE - negatively charged; KHR - positively charged; STNQ - polar; G - glycine; P - proline;	

Tab. 2: Reduced alphabets for predicting CN. Groups are separated by '/'. Solid rectangle marks amino acids that remain in the same group for all four alphabets.

For the three-letter alphabet we can observe a group of our letters, $GHTS$. This group of amino acids surprises the domain experts because it clusters together amino acids having very different properties. G, T and S are small amino acids, H is large. G and T are hydrophobic, while the other two are not. H is aromatic and has a high coil propensity. The generation of this group of amino acids by our protocol is not an artifact. If we go back to the original data with the 20-letter AA alphabet and we check the distribution of CN values in our dataset separately for each amino acid type we can observe that these four amino acids present very similar distributions. Thus, even if originally they have different properties, in relation to CN they behave in a similar way. This issue was successfully captured by our automatic alphabet reduction protocol, and it is a very interesting discovery, because it challenges the preconceptions of the domain experts.

Conclusions

We have applied evolutionary computation tools (for both optimization and machine learning) to tackle a very difficult and relevant domain: Protein Structure Prediction, specifically we have designed a protocol that automatically simplifies the protein representation without losing crucial information, in a process known as alphabet reduction.

The experiments that we have conducted to verify this protocol have shown that our method (1) obtains similar performance to the original AA alphabet, thus achieving the objective of not losing crucial information in the process of reducing the alphabet, (2) obtains better performance than some classic human designed reduced alphabets and (3) the scientific findings obtained by our protocol challenge some of the general understanding of the PSP field. We are currently working on overcoming the problems that we identified in the fitness function of our protocol. We will soon publish improved results.

Acknowledgments

We would like to thank Jonathan D. Hirst for his collaboration in our protein structure prediction research efforts and also Xavier Llorà and Kumara Sastry for the initial work on our alphabet reduction protocol. We acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) under grant GR/T07534/01. We are grateful for the use of the University of Nottingham's High Performance Computer.

Bibliography

[1] Grand challenges 1993: High performance computing and communications, 1992. The FY 1992 U.S. Research and Development Program, Committee on Physical, Mathematical, and Engineering Sciences, Federal Coordinating Council for Science, Engineering, and Technology, Office of Science and Technology Policy.

[2] J. Bacardit. *Pittsburgh Genetics-Based Machine Learning in the Data Mining era: Representations, generalization, and run-time*. PhD thesis, Ramon Llull University, Barcelona, Catalonia, Spain, 2004.

[3] J. Bacardit and N. Krasnogor. Fast rule representation for continuous attributes in genetics-based machine learning. In *GECCO '08: Proceedings of the 10th annual conference on Genetic and evolutionary computation*, page to appear. ACM Press, 2008.

[4] J. Bacardit, M. Stout, J. D. Hirst, K. Sastry, X. Llorà, and N. Krasnogor. Automated alphabet reduction method with evolutionary algorithms for protein structure prediction. In *GECCO '07: Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 346–353, New York, NY, USA, 2007. ACM Press.

[5] J. Bacardit, M. Stout, N. Krasnogor, J. D. Hirst, and J. Blazewicz. Coordination number prediction using learning classifier systems: performance and interpretability. In *GECCO '06: Proceedings of the 8th annual conference on Genetic and evolutionary computation*, pages 247–254. ACM Press, 2006.

[6] B. Broome and M. Hecht. Nature disfavors sequences of alternating polar and non-polar amino acids: implications for amyloidogenesis. *J Mol Biol*, 296(4):961–968, 2000.

[7] J. Cornette, K. Cease, H. Margalit, J. Spouge, J. Berzofsky, and C. DeLisi. Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *J Mol Biol*, 195(3):659–685, 1987.

[8] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & sons, 1991.

[9] G. Escuela, G. Ochoa, and N. Krasnogor. Evolving l-systems to capture protein structure native conformations. In M. Keijzer, A. Tettamanzi, P. Collet, J. van Hemert, and M. Tomassini, editors, *8th European Conference on Genetic Programming*, volume 3447 of *Springer Lecture Notes in Computer Science*, pages 73–83, Lausanne, Switzerland, March, April 2005. EuroGP, Springer Berlin Heidelberg. ISBN 978-3-540-25436-2. Nominated for Best Paper Award.

[10] G. Harik. Linkage learning via probabilistic modeling in the ecga. Technical Report 99010, Illinois Genetic Algorithms Lab, University of Illinois at Urbana-Champaign, 1999.

[11] J. H. Holland and J. S. Reitman. Cognitive systems based on adaptive algorithms. In D. Hayes-Roth and F. Waterman, editors, *Pattern-directed Inference Systems*, pages 313–329. Academic Press, New York, 1978.

[12] N. Krasnogor, B. Blackburne, E. Burke, and J. Hirst. Multimeme algorithms for protein structure prediction. In *Proceedings of the Parallel Problem Solving from Nature VII. Lecture Notes in Computer Science*, volume 2439, pages 769–778, 2002.

[13] N. Krasnogor, W. Hart, J. Smith, and D. Pelta. Protein structure prediction with evolutionary algorithms. In *International Genetic and Evolutionary Computation Conference (GECCO99)*, pages 1569–1601. Morgan Kaufmann, 1999.

[14] N. Krasnogor and D. A. Pelta. *Fuzzy Sets based Heuristics for Optimization*, volume 126 of *Studies in Fuzziness and Soft Computing*, chapter Fuzzy Memes in Multimeme Algorithms: a Fuzzy-Evolutionary Hybrid, pages 49–66. Springer-Verlag Berling Heidelberg New York, 2002.

[15] P. Larranaga and J. Lozano, editors. *Estimation of Distribution Algorithms, A New Tool for Evolutionnary Computation*. Genetic Algorithms and Evolutionnary Computation. Kluwer Academic Publishers, 2002.

[16] Y. Mandel-Gutfreund and L. Gregoret. On the significance of alternating patterns of polar and non-polar residues in beta-strands. *Journal of Molecular Biology*, 323(9):453–461, 2002.

[17] F. Melo and M. Marti-Renom. Accuracy of sequence alignment and fold assessment using reduced amino acid alphabets. *Proteins*, 63:986–995, 2006.

[18] K. M. Misura, D. Chivian, C. A. Rohl, D. E. Kim, and D. Baker. Physically realistic homology models built with rosetta can be more accurate than their templates. *Proc Natl Acad Sci U S A*, 103(14):5361–5366, 2006.

[19] L. R. Murphy, A. Wallqvist, and R. M. Levy. Simplified amino acid alphabets for protein fold recognition and implications for folding. *Protein Eng.*, 13(3):149–152, 2000.

[20] S. Smith. *A Learning System Based on Genetic Algorithms*. PhD thesis, University of Pittsburgh, 1980.

[21] A. D. Solis and S. Rackovsky. Optimized representations and maximal information in proteins. *Proteins*, 38(2):149–164, February 2000.

[22] M. Stout, J. Bacardit, J. D. Hirst, J. Blazewicz, and N. Krasnogor. Prediction of residue exposure and contact number for simplified hp lattice model proteins using learning classifier systems. In *Applied Artificial Intelligence*, pages 601–608, Genova, Italy, 2006. World Scientific.

[23] M. Stout, J. Bacardit, J. D. Hirst, and N. Krasnogor. Prediction of Recursive Convex hull Class Assignments for Protein Residues. *Bioinformatics*, 24(7):916–923, 2008.

[24] M. Stout, J. Bacardit, J. D. Hirst, N. Krasnogor, and J. Blazewicz. From hp lattice models to real proteins: Coordination number prediction using learning classifier systems. In *Applications of Evolutionary Computing, EvoWorkshops 2006*, pages 208–220. Springer LNCS 3907, 2006.

[25] M. Stout, J. Bacardit, J. D. Hirst, R. E. Smith, and N. Krasnogor. Prediction of topological contacts in proteins using learning classifier systems. *Soft Computing, Special Issue on Evolutionary and Metaheuristic-based Data Mining (EMBDM)*, In Press, 2008.

[26] G. Venturini. Sia: A supervised inductive algorithm with genetic search for learning attributes based concepts. In P. B. Brazdil, editor, *Machine Learning: ECML-93 - Proc. of the European Conference on Machine Learning*, pages 280–296. Springer-Verlag, Berlin, Heidelberg, 1993.

[27] J. Wang and W. Wang. A computational approach to simplifying the protein folding alphabet. *Nature Structural Biology*, 6:1033–1038, 1999.

[28] K. Yue, K. M. Fiebig, P. D. Thomas, C. H. Sun, E. I. Shakhnovich, and K. A. Dill. A test of lattice protein folding algorithms. *Proc. Natl. Acad. Sci. USA*, 92:325–329, 1995.

About the authors

Jaume Bacardit received his Ph.D. in 2004 from the Ramon Llull University in Barcelona, Spain. His thesis studied the adaptation of the Pittsburgh approach of Learning Classifier Systems (LCS) to Data Mining tasks. In 2005 he joined the University of Nottingham, UK as a postdoctoral researcher in the ASAP research group under the supervision of N. Krasnogor to work on the application of LCS to data mine large-scale bioinformatics datasets and extract interpretable explanations from the learning process. In 2008 he was appointed as a Lecturer in Bioinformatics at the University of Nottingham. This is a joint post between the schools of Biosciences and Computer Science with the aim of developing interdisciplinary research at the interface of both disciplines. In the School of Computer Science he is part of the ASAP research group. In the School of Biosciences he is part of the Multidisciplinary Centre for Integrative Biology (MyCIB), which hosts the Centre for Plant Integrative Biology (CPIB). His research interests include the application of Learning Classifier Systems and other kinds of Evolutionary Learning to data mine large-scale challenging datasets and, in a general sense, the use of data mining and knowledge discovery for biological domains.

Homepage: http://www.cs.nott.ac.uk/~jqb/
Email: jaume.bacardit @nottingham.ac.uk

Mike Stout received his BSc from the University of Durham in 1983 and MSc in Computing from the University of Essex in 1988. His PhD research in the ASAP research group at the University of Nottingham under the supervision of N. Krasnogor focused on predicting novel aspects of proteins structure with learning classifier systems. Between 1995 and 2002 he headed the Electronic Journals Group at Oxford University Press. Currently, he is currently a Research Fellow in the Center for Plant Integrative Biology at the University of Nottingham. His research interests include Structural Bioinformatics and Evolutionary Learning, in particular Learning Classifier Systems.

Homepage: http://www.cs.nott.ac.uk/~mqs
Email: mqs@cs.nott.ac.uk

Natalio Krasnogor is an Associate Professor within the Automated Scheduling, Optimisation and Planing Research Group (ASAP) within the School of Computer Science at the University of Nottingham. He established and co-chairs the series of international workshops on Memetic Algorithms (WOMA), Nature-Inspired Cooperative Strategies for Optimisation (NICSO), Embodied Evolution (EmboEvo) and the European Conference on Synthetic Biology (ECSB). Dr Krasnogor has published more than 70 refereed papers. He was a guest editor for the Evolutionary Computation Journal, Journal of Fuzzy Sets and Systems, IEEE Transactions on Systems Man and Cybernetics and BioSystems Journal. He is editor of the book Systems Self-Assembly: Multidisciplinary Snapshots to appear in Elseviers Studies in Multidisciplinarity series and edited a book on Memetic Algorithms for Springer. He is associated editor for the Evolutionary Computation journal, Founding Chief-Editor (technical) of the new Springers Journal Memetic Computing and editor for the journal of Simulation and Modelling in Engineering and Journal of Artificial Evolution and Applications. He co-chairs the IEEE Computational Intelligent Society's Emergent Technology Technical Committee on Molecular Computing. He is investigator and co-investigator in grants totalling Ĩ2M from EPSRC, BBSRC and the EU. Krasnogor is a member of the EPSRC peer review college and also a member of Nottinghams Centre for Plant Integrative Biology (www.cpib.eu), one of the 6 EPSRC/BBSRC flagship centres for Systems Biology. He currently supervises 4 Postdoctoral fellows and 9 Ph.D. students.

Homepage: http://www.cs.nott.ac.uk/~nxk
Email: Natalio.Krasnogor@Nottingham.ac.uk

Optimization of Store Performance Using Personalized Pricing

Cem Baydar, Ph.D Director, Peppers & Rogers Group

Currently, most of the grocery stores provide special discounts to their customers under different loyalty card programs. However, since each individual's shopping behavior is not taken into consideration, these discounts do not help optimize the store performance. We believe that a more determined approach such as individual pricing could enable retailers to optimize their store performance by giving special discounts to each customer. Our approach requires each customer is modeled as an agent and his/her shopping behavior is obtained from transaction data. Then, the overall shopping behavior is simulated and the store performance is optimized using Monte-Carlo simulations and evolutionary computation. The results showed that individual pricing outperforms the traditional product-centered approach significantly.

Introduction

As the competition in retail industry increases, retailers are becoming much more obligated to optimize their store performance. Currently, most of the grocery chains in the U.S offer loyalty programs. However, these loyalty programs mostly apply blanket couponing technique by offering the same discounts to their subscribers. However, humans are different and each individual has his/her own preference of products and price levels. Therefore modeling each customer separately and providing him/her individual coupons could improve the store performance. This type of offering is known as one-to-one marketing in the literature. Our proposed approach assumes that by using a sufficiently rich transaction data, it is possible to capture each regular customer's shopping behavior.

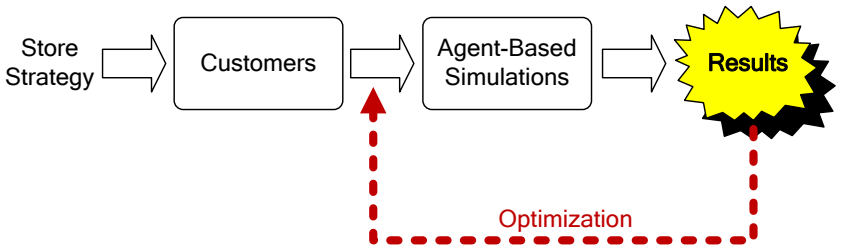


Fig. 1: Outline of the Proposed Approach

Then, individual models (agents) can be generated using this behavioral information and an agent-based system can be developed to simulate overall shopping behavior. The inputs for this agent-based simulation system can be provided by a store manager based on a strategy defined by the relative importance of three factors: profits, sales volume and customer loyalty. Finally, the system can use agent-based simulations in combination with evolutionary computation to identify the set of discounts for each customer. Figure 1 shows the overall approach. We have developed a system and tested the proposed approach against different blanket couponing pricing strategies. The results showed that individual pricing outperforms blanket couponing approach significantly. We believe that retailers can optimize their store performance by applying individual pricing.

Our Approach

One-to-one marketing is a customer relationship management paradigm which aims building customer loyalty by trying to sell as many as products as possible to one customer at a time [2, 3]. Unlike the traditional clustering approach, one-to-one marketing aims to treat each customer as an individual rather than a part of a segment. Grocery retail has always been an interest for the application of one-to-one marketing. In retail industry, most supermarkets use customer loyalty cards and several companies have also started to analyze the premise of one-to-one marketing in addition. The main advantage is that in grocery business almost every customer is a repeated buyer and grocery goods are consumed at a constant rate. Therefore, there is sufficient amount of data to model each regular customer's shopping behavior. Our approach uses an agent-based [1] modeling and simulation approach which is different from the more focused store optimization research approaches found in the literature. In agent-based computational modeling, only equations governing the micro social structure are included (i.e., shopping behavior of each individual). Then, the overall macroscopic structure of the system grows from the bottom-up. Typically for grocery store optimization, revenues, costs and sales volume are taken into account as complex mathematical equations. However in agent-based approach, these values are determined by summing up each customer's shopping activity such as his/her shopping frequency and spending. The implementation steps of our approach are as follows:

1. Model each customer's shopping behavior from transaction data.
2. Create customer models as agents using these models.
3. Perform agent-based simulations and optimize the store performance for a given store strategy.

Problem Statement and Formulation

A grocery store manager has to decide on the store strategy based on the relative importance of three goals: profits, sales volume and customer satisfaction. These goals are contradictory (i.e., a store manager could maximize customer satisfaction by reducing all prices to zero). Therefore, what determines the overall store performance is the difference between each objective. We can visualize the task of setting a store strategy as adjusting the three levers as shown in Figure 2.

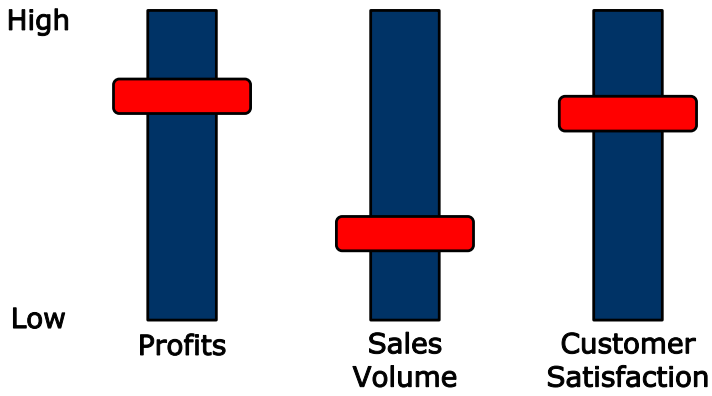


Fig. 2: Three goals to determine store strategy.

The optimization strategy can be defined in mathematical terms as:

$$\text{Maximize} f(x,y,z) = w_1 x + w_2 y + w_3 z \quad (1)$$

where, x is the profit, y is the sales volume, z is the customer satisfaction, while w_1, w_2 and w_3 are the appropriate weights determined by the store manager. Since we are using agent-based models, there is no way of exploring x, y and z dimensions directly. Therefore, they are not the decision variables. The decision variables of this problem are the set of discounted products and discount values for these products. Both of these variables are different for each customer since we are giving individual discounts. Therefore, two questions are being addressed to maximize the objective function:

1. What is the optimal set of products for each customer?
2. What should be the discount values on these products?

Problem Modeling

There are two types of models that we consider for this problem: *store model* and *customer model*.

Store Model. The store model consists of several parameters such as:

- The number of products
- Quantity stored from each product
- Sales price of each product
- Product replenishment frequency
- Replenishment threshold
- Replenishment size
- Daily stock keeping cost of each product (inventory cost)

Customer Model. Each customer is modeled with several shopping properties such as:

- Shopping frequency
- Price sensitivity for each product
- Buying probability for each product
- Consumption rate for each product

Price sensitivity is defined for each product since a customer may have different shopping behavior towards each product. A person's buying probability can be influenced by giving a discount. This change is formulated as:

$$\Delta BP = (1 - \Omega(kd)) \tag{2}$$

where, ΔBP is the change in buying probability, d is the discount rate, k is the price sensitivity, and $\Omega(\cdot)$ is a probabilistic normal distribution function with mean kd and standard deviation $(1/3)kd$. The following formula is used to calculate the updated buying probability:

$$BP(A) = BP'(A)\Delta BP(A) \tag{3}$$

where, $BP(A)$ is the new buying probability of product A after price change, $BP'(A)$ is the buying probability before price change, $\Delta BP(A)$ is the change in buying probability due to the discount offer. In addition to these properties, there are two behavioral rules:

1. As the customer buys a product continuously, he/she starts building loyalty towards that product (i.e., buying probability increases)
2. If the customer finds the prices high for him/her or can not find a product from his/her shopping list, he/she gets frustrated and his/her probability of arrival decreases.

Understanding associations between products is very important when giving individual discounts. For one customer, Pepsi and Coke may be substitutes but for another who likes both products they may be independent. If a discount is given on one of the substitute or complement products, the other product's buying probability will also change. Two types of association are possible between products: *complements* and *substitutes*.

One way of understanding whether two products are dependent is using a statistical dependency test. If two products are independent, then the probability of their co-occurrence is the same as the product of the probabilities of the individual events. For example if Coke and Pepsi occurred separately in 25% of all baskets, then the expected co-occurrence of these two products is 6.25%. Any significant deviance (positive and negative) from this expected value may indicate product dependency.

It is imperative that when giving individual discounts the targeted products should be chosen carefully in order to obtain better store performance. Ineffective discounts may decrease both the customer satisfaction level and profitability. If there are two substitute products A and B, the buying probability of the dependent product B changes according to the given discount on product A using the following formula:

$$\Delta BP(B) = \frac{BP(B)}{BP(A) + BP(B)} - DeltaBP(A) \tag{4}$$

As it can be seen from the equation above, if the change in buying probability of product A is positive, the change in the substitute product is negative. The change is proportional to the relative importance of the buying probabilities between product A and B. For complement products, the change is directly proportional with product A so the negative sign should be removed.

Finally, each customer has a satisfaction function. In order to measure this, we calculate the sum of the buying probabilities of the products which are expected to be purchased by the customer when he/she comes into the store. Then, we calculate the sum of buying probabilities of the products, which were bought in the simulation after discounts. The satisfaction function is defined as the ratio of these two summations as given in the following equation:

$$SF = \frac{\sum BP_a}{\sum BP_e} \tag{5}$$

where, BP_a is the simulated buying probabilities after discounts and BP_e is the expected buying probabilities. As discussed earlier, if a person can not find an item from his/her shopping list or finds the prices high, he/she skips buying that product. Therefore, his/her satisfaction function decreases proportionally depending on the buying probability of that item (i.e., favorite items have much impact on the satisfaction function). This also affects his/her shopping arrival probability.

Optimization

The overall optimization stage is composed of 3 steps:

1. Performing sensitivity analysis on the product space of each customer to select the most suitable products from substitute pairs;
2. Applying the developed optimization algorithm;
3. Ranking of the products to identify the product set for a specified number of discount coupons.

Since discounts should be given on only one product from each substitute group, the first step is reducing the search space by selecting these suitable products. In this step, we pick products one-by-one from each substitute pair and perform sensitivity analysis by applying 1% discount to that product. Then, we simulate the shopping behavior and compare the store performance in profits, sales volume and customer satisfaction between all substitute products. Based on these comparisons, the product which has the most effect on store performance is chosen from each product group. By following this procedure for each customer, we reduce the number of product space for the optimization phase.

In the second step, we apply the optimization algorithm to the set of products selected and obtain the optimal discounts to maximize the store performance. In order to solve this optimization problem, we have developed a hybrid parallel simulated annealing algorithm which uses the survival of the fittest method based on evolutionary computation concepts. At first, the search space is divided into n equal parts and a population of m starting points is selected from each part. Then, using simulated annealing each member starts exploring its neighborhood in a parallel fashion. After each evaluation, better members are replicated while worse members are eliminated from the population based on their fitness value, which is the value of objective function, or in other words, the store strategy.

It should be also noted that we evaluate the objective function $f(S)$, k times using Monte-Carlo simulation since the shopping behavior is probabilistic. This evaluation makes the problem computationally extensive. By eliminating worse members in the population, we also reduce unnecessary computations in a non-promising region and explore a more promising region with multiple members in parallel. Detailed information about this algorithm can be found in our previous work [3].

Case Study

In order to compare the two approaches, we have built a sample database of 200 customers with 100 products from a real grocery store and investigated the performance difference against same allowance on promotion spending. As a promotion strategy, for the following 15 days, we would like to spend $ 1,150 on the discounts and we want to maximize the customer satisfaction.

One possible approach is using a traditional approach such as giving 10% discount on top-10 favorite products. Another approach is by following the individual discounting strategy, giving 10 coupons to each individual at the store entrance with different discount levels on different products. For the optimization process we have selected our objective function as:

$$\text{Maximize} f(x,y,z) = 0.25x + 0.75z \tag{6}$$

where, x represents the profits and z the customer satisfaction. Both approaches were simulated in the developed environment. It was observed that individual pricing outperforms the traditional approach significantly by increasing the customer satisfaction by 8.75%. Figure 3 shows the results.

This and other case studies conducted [4] showed that personalized pricing outperforms the traditional product-centric approach significantly by increasing customer satisfaction and profits. We believe that personalized pricing will again outperform the traditional approach since it optimizes the store performance by looking at each customer's shopping behavior.

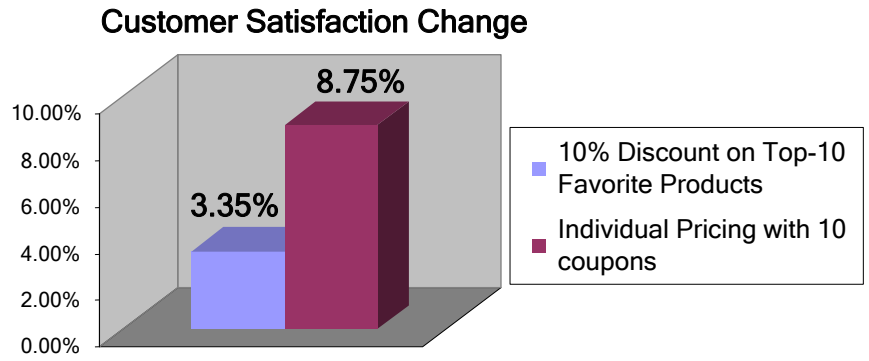


Fig. 3: Results of the case study.

Discussions and Conclusion

For retail sectors with tighter profit margins and where customer loyalty is highly dependent on the prices offered, it is essential to optimize the resources spent on increasing the customer satisfaction. Grocery retail is one of these sectors. Currently, most of the grocery stores provide a type of loyalty program which provides same discounts to subscribed customers. However this product-centered approach is efficient up to some level since customers are being divided into several segments and treated as a part of the segment rather than an individual

Our discussed approach is based on agent-based modeling and simulation, which models each customer's shopping behavior to simulate the store performance. We have developed a system to simulate the shopping behavior and optimize the store performance. We have conducted several case studies using this environment and compared the performance of two approaches. The results showed that individual pricing outperforms the traditional product-centered approach significantly. Several implementations have been conducted with industry partners and encouraging results were achieved. We believe that the discussed approach will impact the grocery retail significantly by increasing the customer satisfaction, sales volume and profits.

Bibliography

[1] Ferber J. (1999), *Evolutionary Computation in Practice*, Addison Wesley.

[2] Peppers D., Rogers M. (1997), *The One to One Future: Building Relationships One Customer at a Time*, Double Day Publications.

[3] Peppers D., Rogers M. (1999), *Enterprise One to One: Tools for Competing in the Interactive Age*, Double Day Publications.

[4] Yu T., Davis L., Baydar C., Roy R., (2002), (2008), *A Hybrid Parallel Simulated Annealing Algorithm to Optimize Store Performance*, Springer, Studies in Computational Intelligence, vol. 88.

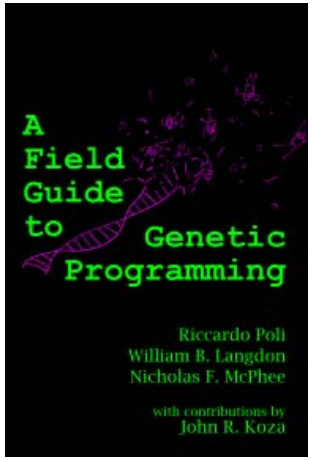
About the author

Cem Baydar is currently working as a Director at Peppers & Rogers Group, a leading strategy and management consulting firm. In this capacity, he worked with many senior executives and helped them craft their marketing and sales strategy to make their corporations more profitable using customer-centric strategies. Prior to joining Peppers & Rogers Group, Cem worked as the Director of Analytical Solutions at comScore Inc., the leading on-line market research and consulting company in the US. Prior to comScore, he worked at Accenture's Innovation Group as Manager for 5 years. Dr. Baydar received his Ph.D from The University of Michigan, Ann Arbor in 2001. With two patents, many published articles and a book; he has a proven track record in innovation, business strategy development, and incubation and evaluation of emerging technologies, including the application of Genetic Algorithms to complex real-world problems.

Company homepage: www.peppersandrogers.com
Email: cem.baydar@gmail.com

Is this the Future of Academic Publishing?

William. B. Langdon, University of Essex, UK



The Field Guide to Genetic Programming [1] has now been freely available as a PDF to down load for three months. According to figures provided by the publisher, lulu.com, during the first three months it was down loaded 11533 times. If the current trend (see Figure 1) continues, the total for the first year would be in the region of 27000 down loads.

While not quite in the same league as Harry Potter, if down loaded copies were equivalent to physical copies, the Field Guide would still be amongst the most successful computer science books. However is it fair to equate something which is delivered at no charge in a few seconds directly to you, with a physical book, which costs real (and in some cases significant amounts of) money and takes days or even weeks to arrive? Obviously not. However from an academic author's perspective, what matters is not what it cost but the impact it has. How many of the people who down load a free PDF will read past the first page? One suspects that the proportion of customers who buy a physical book but never look between the covers, is much lower. There does not seem to be a rapid and reliable way to find out. After several years, books start to show up in citation counts. May be we shall have to wait for these in order to estimate the impact of electronic books.

Despite explicit use of a creative commons license, which explicitly forbids others from laying claim to it or commercially exploiting it, the Field Guide's PDF appeared briefly on a web site which attempted to charge for it. Another, as yet unrealised, fear is that it will be plagiarised. It does not seem possible, even for commercial publishers, to prevent all abuses of Internet resources.

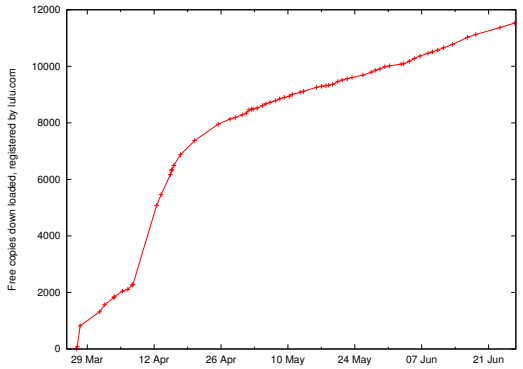


Fig. 1: Copies of "A Field Guide to Genetic Programming" down loaded since its launch at EuroGP on 26 March 2008. There were more than 800 down loads in the first 24 hours. The second steep rise corresponds to the free book being mentioned on a prominent scientific blog in the USA.

According to a very unscientific straw poll, those about to publish books on evolutionary computing, are split. Some still intend to seek contracts with major multinational publishers. And some are intending that their new book will be available as a free electronic down load from the Internet.

The authors' aim, even before writing the book, was that it should be as accessible as possible. Hence the choice of electronic publishing, backed up by a minimal cost print on demand service with rapid postal delivery direct to the reader, from lulu.com, Amazon and Google books, etc. This strategy seems to be working.

Bibliography

[1] Riccardo Poli, William B. Langdon, and Nicholas Freitag McPhee. *A field guide to genetic programming*. Published via lulu.com and freely available at www.gp-field-guide.org.uk, 2008. (With contributions by J. R. Koza).

Genetic Algorithm File Fitter (GAFFitter)

Douglas A. Augusto, daaugusto@gmail.com

Genetic Algorithm File Fitter, *GAFFitter* for short, is a tool based on a genetic algorithm (GA) that tries to fit a collection of items into as few volumes of specific size as possible. For example, the items might be files and the volumes might be CDs or DVDs.

GAFFitter was created with the intent to minimize the number of CDs or DVDs used to store a set of files whose total size is greater than the medium capacity. It was further extended to work directly with any set of items, whether it is composed of files/directories or not.

GAFFitter is characterized by five main features, namely:

- The global search based on a *genetic algorithm*.
- The *filter-oriented* design, that is, a versatile interface suitable for integration with other tools and front-ends.
- The possibility to use filenames as the input and to directly *read a list of items and their sizes*.
- The *great flexibility* provided by the input arguments, which controls the behaviour of GAFFitter, including many genetic algorithm parameters.
- The fact that it is a *free software*, which makes it possible for the users to study, change and redistribute GAFFitter.

The current development version of GAFFitter[1] is based on the Falkenauer's *Hybrid Grouping Genetic Algorithm* (HGGA) [1], which is probably one of the best GA approaches for bin packing problems.

[1] The development version can be fetched from the GAFFitter's Subversion repository (see the website). Be aware, however, that the development versions are usually unstable, non-optimized and prone to bugs.

In the HGGA, the genes represent each one a group of items, i.e., each gene is treated as a bin and their items act as an unit, a building block; therefore, the crossover operator does not mixes items on an individual basis, but, rather, it combines groups of bins. Besides, HGGA uses a local optimizer inspired on the Dominance Criterion of Martello and Toth [1], which basically tries iteratively to replace a few items of a certain bin by fewer items that fit better in. This procedure not only optimizes the bin, but also eases the reallocation of the replaced items, since smaller items are easier to fit.

GAFFitter is written in C++ and is currently available as a command-line program for POSIX-compliant systems (GNU/Linux, BSD derivatives and so on). The simplest way to run GAFFitter is as follows:

```
gaffitter -t 700m *
```

This command will arrange the files and subdirectories of the current directory into sets of at most 700 megabytes (a typical CD), in such a way that the number of sets is minimized. In other words, GAFFitter will try to fit the given files and directories into as few as possible volumes of 700MB.

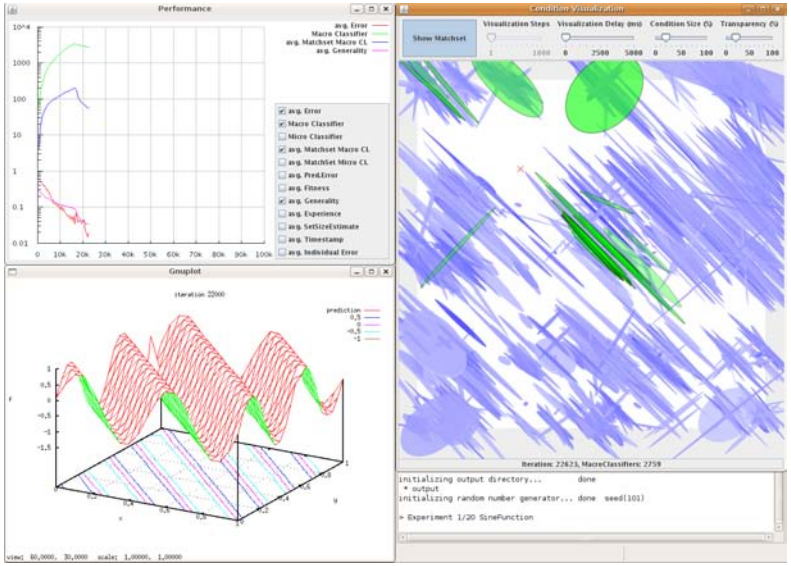
A comprehensive description of GAFFitter's options and parameters, several usage examples, and instructions on how to get its source code can be found on GAFFitter's website at http://gaffitter.sf.net

Bibliography

[1] E. Falkenauer. A Hybrid Grouping Genetic Algorithm for Bin Packing, 1996.

Announcements

XCSF-Ellipsoids Java plus Visualization



From Patrick O. Stalph & Martin V. Butz

XCSF-Ellipsoids Java is an XCSF learning classifier system implementation using hyperellipsoidal conditions and recursive least squares predictions for function approximation. The code can be used to evaluate XCSF on several test functions with online visualization support for performance, prediction, and conditions. Other test functions or approximation problems can be easily implemented. See MEDAL Report No. 2008008 for more information.

www.coboslab.psychologie.uni-wuerzburg.de

medal.cs.umsl.edu/software.php

WCCI 2008 Special Session:

Computational Intelligence on Consumer Games and Graphics Hardware (CIGPU-2008)

Garnett Wilson, Memorial University of Newfoundland, Canada
Simon Harding, Memorial University of Newfoundland, Canada

Overview

Due to its speed, price and availability, there is increasing interest in using mass consumer market commodity hardware for engineering and scientific applications. To date, most of this interest has concentrated upon the highly parallel graphics processing units (GPUs). However, there is increasing interest in using games consoles such as the XBox 360, PlayStation 3 with its the Cell processor, for both research and applications (see gpgpu.org for examples).

The developers of this hardware are well aware that their products can be used for more than simply gaming, and have released a myriad of hardware and software platforms for alternate usage. This new hardware is expected to soon become a widely used technology in scientific computing, and for good reason: the latest GPUs have 256 high-speed, stream processors and are very low cost. Many problems in CI can be implemented using such a platform — and obtain a massive speed increase.

At WCCI 2008, a special session organized by William B. Langdon, Simon Harding, and Man Leung Wong, was held concentrating on how such hardware is beneficial to the computational intelligence (CI) research community. The presenters at this session were all early adopters of this new technology. As such, they must deal with many unknowns. In addition to choosing the right development and hardware path, it is important to think about the type of problems in CI, and the appropriate implementations to best use the hardware. As these are important issues, part of the purpose of the session was to allow for discussion and the exchange of ideas and experiences. In addition to the technical presentations, there were two short panel discussions.

Presentations

Bill Langdon (University of Essex, UK) has filled the need for a fast pseudo-random number generator using Park-Miller. The algorithm was created using Rapidmind with a nVidia GeForce 8800 GTX, but the algorithm was described in detail so as to be generally implementable in other shader or GPGPU languages such Cg, HLSL, Brook, or CUDA. Using C++, the algorithm on GPU was found to be 44 times faster than running Park Miller on the CPU. The code is available via anonymous ftp from cs.ucl.ac.uk genetic/gp-code/random-numbers/gpu park-miller.tar.gz.

Garnett Wilson (Memorial University of Newfoundland, Canada) presented a method for programming the XBox 360 (with execution on both CPU and GPU) to perform linear genetic programming (LGP) using Microsoft's XNA framework. The work, done with Wolfgang Banzhaf, included a number of milestones: it was the first implementation of a research-based GP system on a commercial video game platform, the first implementation of LGP in a GPGPU application, and the first instance of any video game console being used for GPGPU purposes. The presentation examined benchmarking of classification and regression problems in LGP. Fitness evaluation and mutation operations were placed on the CPU and GPU on both Windows and XBox 360 platforms.

Wai-Man Pang (The Chinese University of Hong Kong, Hong Kong) introduced a high-quality pseudo-random number generator (PRNG) by presenting a framework to generate a shader applicable across multiple GPUs. The authors used cellular automata (CA) to yield both high speed and parallel random number generation. The configuration of the CA PRNG was computed automatically by optimizing an objective function that accounts for quality of generated random sequences. Results were inspected by applying the best PRNG shader to photon mapping. Timing statistics showed that the parallelized GPU version of the PRNG was considerably faster than the CPU alternative.

Raghavendra D. Prabhu (National Institute of Technology, Karnataka, India) used Kohonen's Self Organizing Map (SOM) for pattern recognition due to its implicit parallel nature, and with minor modifications was implemented using parallel execution on the GPU. The implementation was constructed with Microsoft Research Accelerator, and experiments indicated that for the largest network sizes considered, use of the GPU provided substantial speedup. The overhead of interleaved use of GPU and CPU was also addressed, as well as design decisions about when to use GPU.

Simon Harding (Memorial University of Newfoundland, Canada) demonstrated the use of GPUs to perform fitness evaluation for evolved image filters. Applying an image filter involves executing the same program on each pixel in an image – a task that is easily parallelisable and fits well with the SIMD architecture of the GPU. The speed advantage of the GPU also allows for multiple images to be tested during fitness evaluation – which is hoped improves the quality of the evolved programs.

Robert Luke (University of Missouri-Columbia, USA) described how to use the parallelism of both the creation of a fuzzy rule base and fuzzy inference to design a GPU-based implementation of a fuzzy logic system. The GPU fuzzy logic implementation outperforms the CPU implementation by over two orders of magnitude under particular configurations, and it outperformed the CPU-side implementation for large data sets.

William B. Langdon (University of Essex, UK) presented another work that used GP to datamine five million correlations between probes within Affymetrix GeneChip HG-U133A probesets across 6685 human tissue samples from NCBI's GEO database. The genetic programming implementation running on a state of the art GPU automatically evolved a biologically feasible predictor of probe quality. Analysis of the resulting population lends support for Affymetrix's claim that poor probe performance is not due to the probe's simple Watson-Crick self-hybridising.

Shinji Fukui (Aichi University of Education, Japan) proposed a method, entirely based on the GPU, to provide a fast and robust means of extracting moving objects from video. The method used a$*$ component and b$*$ components of CIELAB color space to prevent the extraction of the shadow areas to isolate the moving objects. The technique was demonstrated on real video taken in both indoor and outdoor environments.

Panel Discussion

To encourage more interactivity during the session, two short panel discussions were organised. In the first, Simon Harding and Tien-Tsin Wong discussed some issues regarding the technical aspects of implementing computational intelligence techniques using GPUs. For the second, William Langdon and Simon Harding discussed the future of the CIGPU special session.

GPU Programming Issues (Simon Harding)

- Discussed the pros and cons of the various APIs and toolkits available. How to chose a platform?
- GPU programming is still an immature area. One key problem is the current lack of quality documentation. This increases development time and prevents newcomers from using these tools.
- Development tools for GPUs are either basic or targeted towards graphics work. How do we work around such issues?

Shader Programming vs CUDA Issues (Tien-Tsin Wong)

- Comparison of two of the main approaches to GPU programming.
- Shader programming was designed for graphics, and some conventional programming concepts are hard to map to this mindset.
- CUDA is an API for nVidia graphics cards that allows for more general purpose programming methodologies.
- While shader programming is standardised across platforms and manufacturers, CUDA is limtied to nVidia hardware. Does this present problems?

Future of CIGPU (William B. Langdon and Simon Harding)

- There was interest in holding future versions of the CIGPU session.
- It is hoped that there will be greater interest in future events.
- There was enthusiasm for a practical workshop on GPU programming techniques for computational intelligence. This workshop may be a standalone event.
- The possibilities of special issues of journals and a book were discussed.

Summary of the CIGPU Session

The first CIGPU session was a great success, with a number of computational intelligence techniques related to a broad range of areas including bioinformatics, fuzzy logic, image processing and genetic programming - all performed using GPUs or game consoles. Parallel computing and the use of GPUs, and other multi-core processors, is an emerging area that is experiencing rapid growth and quickly becoming a mainstream means of high performance computing. Plans are underway to hold CIGPU 2009, with call for papers and further information to follow.

Material

Links to papers, slides and photographs can be found at:

www.cs.ucl.ac.uk/staff/W.Langdon/cigpu/

More resources on genetic programming on GPUs are located at:

www.gpgpgpu.com

Calls and Calendar

September 2008

New Journal: IEEE Transactions On Computational Intelligence And AI In Games

Homepage: http://ieee-cis.org/pubs/tciaig/ (available soon)

Submissions open August/September 2008

The IEEE TRANSACTIONS ON COMPUTATIONAL INTELLIGENCE AND AI in GAMES (T-CIAIG), published four times a year, publishes archival journal quality original papers in computational intelligence and related areas in artificial intelligence applied to games, including but not limited to video games, mathematical games, human-computer interactions in games, and games involving physical objects. Emphasis will also be placed on the use of these methods to improve performance in and understanding of the dynamics of games, as well as gaining insight into the properties of the methods as applied to games. It will also include using games as a platform for building intelligent embedded agents for the real world. Papers connecting games to all areas of computational intelligence and traditional AI will be considered.

The journal is co-sponsored by the IEEE Computational Intelligence Society, the IEEE Computer Society, the IEEE Consumer Electronics Society and the IEEE Sensors Council. It is technically co-sponsored by the IEEE Systems, Man, and Cybernetics Society, the IEEE Instrumentation and Measurement Society, the IEEE Robotics and Automation Society, and the IEEE Communications Society.

The Journal will begin accepting submissions in August/September 2008, with the first issue to be published in March 2009. The journal is expected to rapidly establish itself as the leading publication in the field, following in the tradition of many other high quality IEEE Transactions.

For more information contact the Editor in Chief, Simon M. Lucas, University of Essex, UK, sml@essex.ac.uk.

PPSN 2008 - Parallel Problem Solving from Nature

September 13-17, 2008, Dortmund, Germany

Homepage: http://www.ppsn2008.org/

Call for paper: download

PPSN X will showcase a wide range of topics in Natural Computing including, but not restricted to: Evolutionary Computation, Quantum Computation, Molecular Computation, Neural Computation, Artificial Life, Swarm Intelligence, Artificial Ant Systems, Artificial Immune Systems, Self-Organizing Systems, Emergent Behaviors, and Applications to Real-World Problems.

Paper Presentation

Following the now well-established tradition of PPSN conferences, all accepted papers will be presented during small poster sessions of about 16 papers. Each session will contain papers from a wide variety of topics, and will begin by a plenary quick overview of all papers in that session by a major researcher in the field. Past experiences have shown that such presentation format led to more interactions between participants and to a deeper understanding of the papers. All accepted papers will be published in the Proceedings.

Sixth International Conference on Ant Colony Optimization and Swarm Intelligence

September 22–24, 2008. Brussels, Belgium

Homepage: http://iridia.ulb.ac.be/ants2008/

Swarm intelligence is a relatively new discipline that deals with the study of self-organizing processes both in nature and in artificial systems. Researchers in ethology and animal behavior have proposed many models to explain interesting aspects of social insect behavior such as self-organization and shape-formation. Recently, algorithms inspired by these models have been proposed to solve difficult computational problems.

An example of a particularly successful research direction in swarm intelligence is **ant colony optimization**, the main focus of which is on discrete optimization problems. Ant colony optimization has been applied successfully to a large number of difficult discrete optimization problems including the traveling salesman problem, the quadratic assignment problem, scheduling, vehicle routing, etc., as well as to routing in telecommunication networks. Another interesting approach is that of **particle swarm optimization**, that focuses on continuous optimization problems. Here too, a number of successful applications can be found in the recent literature. [...]

ANTS 2008 will give researchers in swarm intelligence the opportunity to meet, to present their latest research, and to discuss current developments and applications.

The three-day conference will be held in Brussels, Belgium, on September 22–24, 2008. Tutorial sessions will be held in the mornings before the conference program.

Further Information

Up-to-date information will be published on the web site http://iridia.ulb.ac.be/ants2008/. For information about local arrangements, registration forms, etc., please refer to the above-mentioned web site or contact the local organizers at the address below.

Conference Address

ANTS 2008
IRIDIA CP 194/6 — Tel +32-2-6502729
Université Libre de Bruxelles — Fax +32-2-6502715
Av. F. D. Roosevelt 50 — http://iridia.ulb.ac.be/ants2008
1050 Bruxelles, Belgium — email: ants@iridia.ulb.ac.be

ICES 2008 - 8th International Conference of Evolvable Systems: From Biology to Hardware

September 21-24, 2008. Prague, Czech Republic
Homepage: http://www.fit.vutbr.cz/events/ices2008

The 8th International Conference of Evolvable Systems (ICSE 2008) which will be held in Prague, September 21-24, 2008. Topics to be covered include, but are not limited to: Evolutionary hardware design Evolutionary circuit diagnostics and testing, Self-reconfiguring/repairing and fault tolerant systems, co-evolution of hybrid systems, generative and developmental approaches, embryonic hardware, hardware/software co-evolution, intrinsic and extrinsic evolution, real-world applications of evolvable hardware, on-line hardware evolution, MEMS and nanotechnology in evolvable hardware, evolutionary robotics, formal models for bio-inspired hardware systems adaptive computing, novel devices/testbeds/tools for evolvable hardware.

December 2008

CIG 2008 - IEEE Symposium on Computational Intelligence and Games

December 15-18, 2008, Perth, Australia
Homepage: http://www.csse.uwa.edu.au/cig08/
Deadline August 15, 2008
Submission website: [WWW]

Games have proven to be an ideal domain for the study of computational intelligence as not only are they fun to play and interesting to observe, but they provide competitive and dynamic environments that model many real-world problems. This symposium, sponsored by the IEEE Computational Intelligence Society with technical co-sponsorship from the IEEE Consumer Electronics Society, aims to bring together leading researchers and practitioners from both academia and industry to discuss recent advances and explore future directions in this field.

- Learning in games
- Coevolution in games
- Neural-based approaches for games
- Fuzzy-based approaches for games
- Opponent modelling in games
- Theoretical or empirical analysis of computational intelligence techniques for games
- Comparative studies (e.g. evolved players versus human-designed players or other learning algorithms)
- Multi-agent and multi-strategy learning
- Applications of game theory
- Board and card games

- Economic or mathematical games
- Imperfect information and non-deterministic games
- Evasion (predator/prey) games
- Console and video games
- Realistic games for simulation or training purposes
- Player satisfaction in games
- Games for mobile or digital platforms
- Games involving control of physical objects
- Games involving physical simulation

The symposium will consist of a single track of oral presentations, tutorial and workshop/special sessions, and live competitions. The proceedings will be published by the IEEE and made freely available on this website after the symposium.

The paper submission is now open.

January 2009

FOGA X - Foundations of Genetic Algorithms

January 9-11, 2009, Orlando, Florida USA

Homepage: http://www.sigevo.org/foga-2009

Deadline September 1, 2008

We invite submissions of extended abstracts for the tenth Foundations of Genetic Algorithms, sponsored by ACM SIGEVO. FOGA focuses on theoretical foundations of evolutionary computation. It will be held from Friday, January 9 until Sunday January 11, 2009 in Orlando, Florida in the USA. Attendance will be limited to individuals who have submitted papers, or those requesting attendance in advance. Students are particularly encouraged to participate. Submissions should address theoretical issues in evolutionary computation. Papers that consider foundational issues, place analysis in the context of the wider community of theoretical computer science, or focus on bridging the gap between theory and practice are particularly encouraged.

These topics do not preclude the acceptance of papers that use an experimental approach, but such work should be directed toward validation of suitable hypotheses concerning foundational matters.

Extended abstracts should be between 10-12 pages long. To submit, please email a compressed postscript or a PDF file to foga09@ist.ucf.edu no later than Monday, August 18, 2008. In their submission message, authors should provide the title of the paper, as well as the name, address and affiliation of the author(s).

Submitted papers should use standard spacing and margins, with 11pt or 12pt font for the main text. Authors using LATEX should either use the standard article style file or the FOGA style file which can be found at the conference web-site. A double-blind reviewing process will be employed, so authors are asked to remove references to themselves from their paper. Notification will be in early November, and drafts of the full paper will be due one month after notification. These drafts will be distributed as part of a preprint to participants at FOGA. Authors of papers presented at the FOGA workshop will be asked to contribute final versions of their papers (based on discussion/feedback at the meeting) as part of the final volume.

Further Information

Extended abstracts due	August 18, 2008
Requests for attendance due	September 1, 2008
Notification of acceptance	early November, 2008
Full papers due	early December, 2008
FOGA Workshop	January 9-11, 2009

Organizing Committee

Thomas Jansen,	Thomas.Jansen@tu-dortmund.de
Ivan Garibay,	igaribay@cs.ucf.edu
R. Paul Wiegand,	wiegand@ist.ucf.edu
Annie S. Wu,	aswu@cs.ucf.edu

Further Information

Enquiries and submissions: foga09@ist.ucf.edu

March 2009

2009 IEEE Symposium Series on Computational Intelligence
March 30 - April 2, 2009, Nashville, TN, USA
Homepage: www.ieee-ssci.org
Deadline October 31, 2009

This international event promotes all aspects of the theory and applications of computational intelligence, by hosting 24 technical meetings in one location. Sponsored by the IEEE Computational Intelligence Society, this event will attract top researchers, professionals, and students from around the world. The registration, which will allow participants to attend any session in any technical meeting, will also include the complete set of the proceedings of all the meetings, coffee breaks, lunches, and the banquet. The event will be held in the magical town of Nashville, city of the country music.

Symposia And Workshops

- IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL 2009)
- IEEE Symposium on Computational Intelligence in Control and Automation (CICA 2009)
- IEEE Workshop on Robotic Intelligence in Informationally Structured Space (RiiSS 2009)
- IEEE Workshop on Evolving and Self-Developing Intelligent Systems (ESDIS 2009)
- IEEE Workshop on Evolvable and Adaptive Hardware (WEAH 2009)
- IEEE Swarm Intelligence Symposium (SIS 2009)
- IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB 2009)
- IEEE Symposium on Artificial Life (ALIFE 2009)
- IEEE Symposium on Multicriteria Decision-Making (MCDM 2009)
- IEEE Symposium on Computational Intelligence in Scheduling (CI-Sched 2009)
- IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2009)
- IEEE Workshop on Hybrid Intelligent Systems (HIS 2009)
- IEEE Workshop on Organic Computing (OC 2009)
- IEEE Symposium on Intelligent Agents (IA 2009)
- IEEE Workshop on Memetic Algorithms (WOMA 2009)
- IEEE Symposium on Computational Intelligence in Cyber Security (CICS 2009)
- IEEE Symposium on Computational Intelligence in Vehicles and Vehicular Systems (CIVVS 2009)
- IEEE Workshop on Computational Intelligence in Aerospace Applications (CIAA 2009)
- IEEE Symposium on Computational Intelligence for Multimedia Signal and Vision Processing (CIMSVP 2009)
- IEEE Symposium on Computational Intelligence for Image Processing (CIIP 2009)
- IEEE Workshop on Computational Intelligence in Virtual Environments (CIVE 2009)
- IEEE Workshop on Computational Intelligence for Visual Intelligence (CIVI 2009)
- IEEE Workshop on Computational Intelligence in Biometrics: Theory, Algorithms, and Applications (CIB 2009)
- IEEE Symposium on Computational Intelligence for Financial Engineering (CIFEr 2009)

Important Dates

- Paper Submission Due: October 31, 2008
- Notification to Authors: November 30, 2008
- Camera-Ready Papers Due: January 15, 2009

Please visit www.ieee-ssci.org for call for papers, guidelines, submission information, additional details, and up to the minute information.

2009 EVO*
www.evostar.org
tübingen::germany

April 2009

Evostar 2009 - EuroGP, EvoCOP, EvoBIO and EvoWorkshops
April 15-17, 2009, Tübingen, Germany
Homepage: www.evostar.org
Important Dates for all events are:
Deadline November 5, 2008

The EuroGP, EvoCOP and EvoBIO conferences and the workshops collectively entitled EvoWorkshops compose EVO*: Europe's premier co-located events in the field of Evolutionary Computing. Featuring the latest in theoretical and applied research, EVO* topics include recent genetic programming challenges, evolutionary and other meta-heuristic approaches for combinatorial optimisation, evolutionary algorithms, machine learning and data mining techniques in the biosciences, in numerical optimisation, in music and art domains, in image analysis and signal processing, in hardware optimisation and in a wide range of applications to scientific, industrial, financial and other real-world problems.

EuroGP

Twelfth European Conference on Genetic Programming: high quality papers are sought on topics strongly related to the evolution of computer programs, ranging from theoretical work to innovative applications.

EvoCOP

Ninth European Conference on Evolutionary Computation in Combinatorial Optimisation: practical and theoretical contributions are invited, related to evolutionary computation techniques and other meta-heuristics for solving combinatorial optimisation problems.

EvoBIO

Seventh European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics: the emphasis is on evolutionary computation and other advanced techniques addressing important problems in molecular biology, proteomics, genomics and genetics, that have been implemented and tested in simulations and on real-life datasets.

EvoWorkshops

The twelve workshops which make up this event are focused on the use of Evolutionary Computation in different application areas:

- EvoCOMNET: Telecommunication networks and other parallel and distributed systems
- EvoENVIRONMENT: Environmental issues
- EvoFIN: Finance and economics
- EvoGAMES: Games
- EvoHOT: Design automation
- EvoIASP: Image analysis and signal processing
- EvoINTERACTION: Interactive evolution and humanized computational intelligence
- EvoMUSART: Music, sound, art and design
- EvoNUM: Continuous parameter optimisation
- EvoPHD: Graduate student workshop on evolutionary computation
- EvoSTOC: Stochastic and dynamic environments
- EvoTRANSLOG: Transportation and logistics

In 2009, the event will take place in Tübingen, a traditional university town in Baden-Württemberg, Germany, situated on a ridge between the Neckar and Ammer rivers in the southwest of the country, about 30 kms southwest of Stuttgart. EVO* 2009 will be hosted at Eberhard Karls University in Tübingen, founded in 1477 and one of the oldest universities in Germany.

EVO* 2009 proceedings will be published by Springer Verlag in the Lecture Notes in Computer Science series.

The website www.evostar.org offers information relevant to all events, including calls for papers, deadlines, organising committees, submission requirements, local information and a thorough view on the previous editions.

May 2009

1st International Symposium
on Search Based Software Engineering

May 13-15, 2009, Cumberland Lodge, Windsor, UK

Homepage: www.ssbse.org

Deadline December 5, 2008

We are pleased to announce SSBSE 2009, the inaugural meeting of an annual symposium dedicated to Search Based Software Engineering (SBSE) held in cooperation with the IEEE. The symposium's objective is to build on the recent flourishing of interest in SBSE by not only creating a welcoming forum for discussion and dissemination, but also by establishing a regular event that will strengthen the rapidly growing international community.

The venue for SSBSE 2009 is Cumberland Lodge, a beautiful and historic royal residence in heart of Windsor Great Park with world-class conference facilities. The Lodge is close to London, and only a short taxi ride from Heathrow Airport, yet is surrounded by some of the finest parkland in the country.

The symposium program includes three keynote speakers who are internationally renowned leaders in their research fields:

- Enrique Alba, University of Malaga, Spain
- Lionel C. Briand, Simula Research Lab & University of Oslo, Norway
- David E. Goldberg, Illinois Genetic Algorithms Laboratory (IlliGAL),
- University of Illinois at Urbana-Champaign, USA

We invite the submission of high quality papers describing original and significant work in all aspects of Search Based Software Engineering, including theoretical work, research on SBSE applications, empirical studies, and reports from industrial experience. Applications may be drawn from throughout the software engineering lifecycle. Search methods may include, but are not limited to, operational research techniques and optimisation methods inspired by nature, such as evolutionary algorithms and simulated annealing. We particularly encourage papers that use novel search techniques and describe software engineering applications to which SBSE has not previously been applied.

We also invite papers for a special PhD student track that will provide a venue for students to showcase their SBSE research and to receive feedback from senior members of the SBSE research community. Papers submitted to this track should be no more than 4 pages using the regular conference format. Each paper will be reviewed by selected members of the program committee. To be eligible a student must be registered on a doctoral programme and must not yet have made their final dissertation defence. It is the expectation that the student will give the presentation at the symposium, though there may be other authors on the paper.

Papers must not have been previously published nor have been submitted to, or be in consideration for, any journal, book, or other conference. Papers will be evaluated by members of the program committee based on their originality, technical soundness and quality of presentation. Papers must conform to the IEEE proceedings paper format guidelines, and must be limited to 10 (ten) pages. In addition to full-length papers, we also welcome shorter (4 pages or less) position papers describing novel ideas not yet fully developed or validated. If the paper is accepted, at least one author is expected to attend the conference and to present the paper. Accepted papers will appear in proceedings published by the IEEE CS Press. Further information on paper formatting and submission is available from the symposium website, http://www.ssbse.org

The authors of selected papers that include significant experimental work will be invited to submit extended versions of their papers for a special issue of Empirical Software Engineering journal (EMSE) edited by Springer.

Key Dates

- Paper submission by: 5 December 2008
- Acceptance notification: 27 February 2009
- Camera-ready paper submission by: 13 March 2009

Organizing Committee

- General Chair: Mark Harman, King's College London, UK
- Program Co-Chairs: Massimiliano Di Penta (University of Sannio, Italy) and Simon Poulding (University of York, UK)
- PhD Student Track Chair: Myra B. Cohen, University of Nebraska, Lincoln, USA
- Publicity Chair: Per Kristian Lehre, University of Birmingham, UK
- Website: Paul Emberson, University of York, UK
- Sponsors: Berner & Mattner, Germany; Engineering and Physical Science Research Council (EPSRC), UK

Event organised in cooperation with the IEEE

2009 IEEE Congress on Evolutionary Computation (CEC 2009)

May 18-21, 2009, Trondheim, NORWAY

Homepage: www.cec-2009.org

Deadline November 1, 2008

The 2009 IEEE Congress on Evolutionary Computation (CEC 2009) will be at the Nova Conference Centre and Cinema, Trondheim, Norway during May 18-21, 2009. Sponsored by the IEEE Computational Intelligence Society, co-sponsored by the Evolutionary Programming Society and the Institution of Engineering and Technology, CEC 2009 continues the successful sequence of World-class events going back to 1999.

CEC 2009 will feature a world-class conference that will bring together researchers and practitioners in the field of evolutionary computation and computational intelligence from all around the globe. Technical exchanges within the research community will encompass keynote speeches, special sessions, tutorials, panel discussions as well as poster presentations. On top of these, participants will be treated to a series of social functions, receptions and networking sessions, which will serve as a vital channel to establish new connections and foster everlasting friendship among fellow researchers. The annual IEEE Congress on Evolutionary Computation (CEC) is one of the leading events in the area of evolutionary computation.

CEC covers all topics in evolutionary computation, including, but not limited to: Ant colony optimization, Artificial immune systems, Artificial life, Autonomous mental & behaviour development, Bioinformatics & bioengineering, Coevolution & collective behaviour, Cognitive systems & applications, Combinatorial & numerical optimization, Computational finance & economics, Constraint & uncertainty handling, Estimation of distribution algorithms, Evolutionary data mining, Evolutionary design, Evolutionary games, Evolvable hardware & software, Evolutionary intelligent agents, Evolutionary learning systems, Evolving neural networks & fuzzy systems, Molecular & quantum computing, Particle swarm intelligence, Representation & operators, etc.

Researchers are invited to contribute high-quality papers to CEC 2009. All papers are to be submitted electronically through the Congress website by November 1, 2008. All submitted papers will be refereed by experts in the fields based on the criteria of originality, significance, quality, and clarity. In addition, we are looking for high quality proposals for Special Sessions and Tutorials for the Congress. More details on all of all of these are below.

Call for Contributed Papers

Prospective authors are invited to contribute high-quality papers to CEC2009. All papers are to be submitted electronically through the Congress website. For general inquiries, please contact General Chair Andy Tyrrell at amt@ohm.york.ac.uk. For program inquiries, contact Program Chair Pauline Haddow at Pauline.Haddow@idi.ntnu.no.

Call for Special Sessions

CEC 2009 Program Committee solicits proposals for special sessions within the technical scopes of the congress. Special sessions, to be organised by internationally recognised experts, aim to bring together researchers in special focused topics. Papers submitted for special sessions are to be peer-reviewed with the same criteria used for the contributed papers. Researchers interested in organising special sessions are invited to submit formal proposals to the Special Session Chair Jon Timmis at jt517@ohm.york.ac.uk. A special session proposal should include the session title, a brief description of the scope and motivation, names, contact information and brief CV of the organisers.

Call for Tutorials

CEC 2009 will also feature pre-congress tutorials covering fundamental and advanced computational intelligence topics. A tutorial proposal should include title, outline, expected enrollment and presenter biography. Tutorials are expected to run for 2 hours each. Researchers interested in organising tutorials are invited to submit formal proposals to the Tutorial Chair Stephen Smith at sls@ohm.york.ac.uk.

Important Dates:

- Special Session proposals: September 1, 2008
- Paper submissions: November 1, 2008
- Tutorial proposals: December 1, 2008
- Notification of acceptance: January 16, 2009
- Final paper submission: February 16, 2009

More information can be found at: www.cec-2009.org.

For general inquiries, please contact General Chair Andy Tyrrell at amt@ohm.york.ac.uk.

June 2009

2009 World Summit on Genetic and Evolutionary Computation
June 12-14, 2009, Shanghai, China
Homepage: http://www.sigevo.org/gec-summit-2009
Deadline December 6, 2008
Author notification: January 30, 2009
Camera-ready: March 1, 2009
Author Registration Deadline: March 23, 2009
Conference Dates: June 12-14, 2009

The 2009 World Summit on Genetic and Evolutionary Computation (2009 GEC Summit) will be held June 12-14, 2009, in Shanghai, China. It is sponsored and organized by ACM/SIGEVO, the Special Interest Group for Genetic and Evolutionary Computation, sponsor of the annual GECCO conferences, and will feature the latest research and demonstrated successes in this dynamic area, including new approaches and breakthrough applications to problems in biology, medicine, engineering design, agriculture, logistics, traffic, security, scheduling, military affairs, and other fields.

Topics (including, but not limited to): Genetic Algorithms, Genetic Programming, Evolution Strategies, Evolutionary Programming Coevolution, Learning Classifier Systems, Ant Colony Optimization, Swarm Intelligence/Particle Swarm, DNA-Based Evolutionary Computation, Interactive Computational Models, Evolutionary Multiobjective Optimization, Evolutionary Combinatorial Optimization, Evolutionary Scheduling and Routing, Simulated Annealing, Estimation of Distribution Algorithms, Tabu Search, Biological Applications, Medical Applications, Industry Applications, Agricultural Applications, Military and Security Applications, Artificial Life, Search-Based Software Engineering, Evolutionary Robotics, Nature-Inspired Computation, Swarm Intelligence Optimization Applications, Intelligent Control, Intelligent Management, Intelligent Information Processing, Multi-Agent Theory, Pattern Recognition, Web Intelligence, Intelligent Transportation Systems, and others.

Organizers

Executive Chairs:	Qidi Wu Tongji Univ., China
	David Goldberg, University of Illinois, USA
Advisory Committee Chairs	Ruwei Dai, Inst. of Automation, Chinese Academy of Science, China
	John Koza, Stanford Univ., USA
Co-Chair	Dongyuan Yang, Tongji Univ., China
General Chairs	
	Lihong Xu, Tongji Univ., China
	Erik D. Goodman, Michigan State Univ., USA
Program Committee Chairs	
	Guoliang Chen, Univ. of Sci. & Tech. of China
	Darrell Whitley, Colorado State Univ., USA
Co-Chair	Yongsheng Ding Donghua Univ., China
Local Arrangements Chairs	
	Wanggen Wan Shanghai Univ., China
	Mark Kotanchek, Evolved Analytics, USA
Co-Chair	Xiaoguang Yang Tongji Univ., China
Treasurer	Jie Chen Tongji Univ., China

Submission of Papers

Authors are invited to submit a paper by December 6, 2008. Papers must not exceed 8 pages, must be formatted using the ACM template, and must meet all ACM formatting requirements. All paper submission will be in electronic form only. The template and detailed directions for paper formatting and submission will be found at the conference web site, www.sigevo.org/gec-summit-2009

Review Process

There is a *blind* review process for all submitted papers, with reviews conducted by the Program Committee and other qualified reviewers under their supervision. Acceptance is based on content, presentation and suitability for the conference. Depending on submission volume, some papers may be accepted as poster papers, rather than scheduled for individual presentation times. In that case, the full paper will still appear in the conference proceedings CD, but the author will prepare a poster and present in a special poster session. At least one author of each accepted paper must register for the conference no later than March 23, 2009, or the paper will be withdrawn from the proceedings.

Publication

All accepted papers will be published on CD and in the ACM Digital Library, which is indexed by SCI and EI. Some selected papers will be recommended for publication in expanded versions in such journals as Evolutionary Computation or Genetic Programming and Evolvable Machines, indexed by SCI. All registered participants will receive a CD containing the proceedings.

Presentation of Papers

All accepted papers will be appeared in the proceedings CD and be presented at the conference. Papers accepted for oral presentation will be given 25-minutes to present. LCD projectors will be provided, but each presenter must bring/arrange for a laptop computer to display the presentation (i.e., PowerPoint). Easels will be provided for each accepted poster.

GEC Summit is sponsored by the Association for Computing Machinery Special Interest Group for Genetic and Evolutionary Computation.

July 2009

GECCO 2009 - Genetic and Evolutionary Computation Conference

July 8-12, 2009, Montréal, Canada

Homepage: http://www.sigevo.org/gecco-2009

Deadline January 14, 2009

Author notification: March 11, 2009

Camera-ready: April 22, 2009

The Genetic and Evolutionary Computation Conference (GECCO-2009) will present the latest high-quality results in the growing field of genetic and evolutionary computation.

Topics include: genetic algorithms, genetic programming, evolution strategies, evolutionary programming, real-world applications, learning classifier systems and other genetics-based machine learning, evolvable hardware, artificial life, adaptive behavior, ant colony optimization, swarm intelligence, biological applications, evolutionary robotics, coevolution, artificial immune systems, and more.

Organizers

General Chair:	Franz Rothlauf
Editor-in-Chief:	Günther Raidl
Business Committee:	Wolfgang Banzhaf
	Erik Goodman
	Una-May O'Reilly
Publicity Chair:	Martin Pelikan
Workshops Chair:	Anna I. Esparcia

Competitions Chairs:	Pier Luca Lanzi
Tutorials Chair:	Martin V. Butz
Late Breaking Papers Chair:	TBA
Local Chair:	Christian Gagné
EC in Practice Chairs:	David Davis
	Jörn Mehnen
Graduate Student Workshop Chair:	Steve Gustafson
Undergraduate Student Workshop Chair:	Frank Moore
	Clare Bates Congdon
	Larry Merkle

Important Dates

Paper Submission Deadline	January 14, 2009
Decision Notification	March 11, 2009
Camera-ready Submission	April 22, 2009

Venue

Delta Centre-Ville hotel is located in the heart of downtown, where Old Montreal and new Montreal blend seamlessly, and adjacent to vibrant nightlife, boutique shops and eclectic cuisine. For more information on Delta Centre-Ville, please visit:

www.deltahotels.com/hotels/hotels.php?hotelId=35

Visiting GECCO-2009 will be a great opportunity to visit the famous Montreal Jazz Festival (July 2-12, 2009):

www.montrealjazzfest.com/Fijm2008/festival_en.aspx

More Information

Visit www.sigevo.org/gecco-2009 for information about electronic submission procedures, formatting details, student travel grants, the latest list of tutorials and workshop, late-breaking papers, and more.

For technical matters, contact Conference Chair Franz Rothlauf at rothlauf@uni-mainz.de. For conference administration matters contact Primary Support Staff at gecco-admin@tigerscience.com.

GECCO is sponsored by the Association for Computing Machinery Special Interest Group for Genetic and Evolutionary Computation.

About the Newsletter

SIGEVOlution is the newsletter of SIGEVO, the ACM Special Interest Group on Genetic and Evolutionary Computation.

To join SIGEVO, please follow this link [WWW]

Contributing to SIGEVOlution

We solicit contributions in the following categories:

Art: Are you working with Evolutionary Art? We are always looking for nice evolutionary art for the cover page of the newsletter.

Short surveys and position papers: We invite short surveys and position papers in EC and EC related areas. We are also interested in applications of EC technologies that have solved interesting and important problems.

Software: Are you are a developer of an EC software and you wish to tell us about it? Then, send us a short summary or a short tutorial of your software.

Lost Gems: Did you read an interesting EC paper that, in your opinion, did not receive enough attention or should be rediscovered? Then send us a page about it.

Dissertations: We invite short summaries, around a page, of theses in EC-related areas that have been recently discussed and are available online.

Meetings Reports: Did you participate to an interesting EC-related event? Would you be willing to tell us about it? Then, send us a short summary, around half a page, about the event.

Forthcoming Events: If you have an EC event you wish to announce, this is the place.

News and Announcements: Is there anything you wish to announce? This is the place.

Letters: If you want to ask or to say something to SIGEVO members, please write us a letter!

Suggestions: If you have a suggestion about how to improve the newsletter, please send us an email.

Contributions will be reviewed by members of the newsletter board.

We accept contributions in LaTeX, MS Word, and plain text.

Enquiries about submissions and contributions can be emailed to editor@sigevolution.org.

All the issues of SIGEVOlution are also available online at www.sigevolution.org.

Notice to Contributing Authors to SIG Newsletters

By submitting your article for distribution in the Special Interest Group publication, you hereby grant to ACM the following non-exclusive, perpetual, worldwide rights:

- to publish in print on condition of acceptance by the editor
- to digitize and post your article in the electronic version of this publication
- to include the article in the ACM Digital Library
- to allow users to copy and distribute the article for noncommercial, educational or research purposes

However, as a contributing author, you retain copyright to your article and ACM will make every effort to refer requests for commercial use directly to you.