

# Introdução à Computação Biogeográfica

Rodrigo Pasti, Fernando J. Von Zuben (Orientador), Leandro N. de Castro\* (Co-orientador)

Departamento de Engenharia de Computação e Automação Industrial (DCA)  
Faculdade de Engenharia Elétrica e de Computação (FEEC)  
Universidade Estadual de Campinas (Unicamp)  
Campinas, SP, Brasil

\*Laboratório de Computação Natural (LCoN), Faculdade de Computação e Informática e Programa de Pós-Graduação em Engenharia Elétrica (PPGEE), Universidade Presbiteriana Mackenzie  
São Paulo, SP, Brasil

{rpasti, vonzuben}@dca.fee.unicamp.br, lnunes@mackenzie.br

**Abstract** – In this paper, we will briefly introduce the Biogeographic Computation. This is an initiative to formalize a conceptual framework dedicated to the study and abstraction of processes defined from the observation of nature through the science of biogeography, aiming at the proposition of a new paradigm of optimization and search algorithms, in which time and space have a core role in the conduction of the involved processes. With the computational implementation of concepts such as speciation, gene flow and habitat, the proposal of population-based metaheuristics is achieved, characterized by the promotion of processes which are generally absent, when considered together, or which play a subordinate role in alternative population meta-heuristics: (i) automatic control of the population size; (ii) automatic control of the number of species in the population; (iii) diversity maintenance in the population; and (iv) multimodal search capability.

**Keywords** – biogeography, population metaheuristics, multimodal optimization in continuous spaces.

## 1. Por que a Biogeografia?

Uma parte dos esforços em se desenvolver novas ferramentas computacionais para a solução de problemas complexos é no sentido de entender mecanismos, comportamentos e dinâmicas de sistemas naturais (químicos, físicos e biológicos) e implementar algoritmos inspirados ou baseados nestas ideias. A linha de pesquisa que emprega esta filosofia pode ser denominada de Computação Natural [1], mais especificamente sua sub-área denominada de Computação Inspirada na Biologia ou Computação Bioinspirada [1][2]. Esta sub-área consiste no estudo de fenômenos naturais, comportamentos de organismos vivos e modelos teóricos de sistemas naturais de modo a criar ferramentas computacionais (meta-heurísticas) inspiradas nestes sistemas. Neste contexto, estão inseridas as Redes Neurais Artificiais [3], a Computação Evolutiva [4] e [5], os Sistemas Imunológicos Artificiais [6] e a Inteligência de Enxame [7].

Em todas as áreas citadas, é possível encontrar mecanismos, definições e conceitos básicos e avançados que constituem suas respectivas estruturas conceituais (framework). Através dessas estruturas conceituais, é possível projetar algoritmos de diversos tipos para a resolução de diferentes classes de problemas. Embora nem sempre essas estruturas sejam formalizadas e unificadas, é possível encontrar esforços nesse sentido, por exemplo, a

Engenharia Imunológica [6] como uma estrutura conceitual formal para os Sistemas Imunológicos Artificiais; e os volumes [4] e [5] que apresentam os elementos básicos de projeto de algoritmos para a Computação Evolutiva.

Com o objetivo de abstrair conceitos, definições e mecanismos, este trabalho propõe uma estrutura conceitual básica baseada em processos presentes na ciência da Biogeografia. Antes de descrever qualquer teoria, entretanto, é preciso fazer a seguinte pergunta: dentre tantas bioinspirações distintas e tantas técnicas resultantes, por que escolher a Biogeografia? Para iniciar a resposta a essa pergunta, considere a seguinte frase, extraída de [8]: “Os seres vivos são extremamente multiformes. Há provavelmente algo entre 5 e 50 milhões de espécies, menos de 2 milhões foram formalmente reconhecidas e estão catalogadas. Milhões ou talvez bilhões de espécies viveram em algum lugar do passado, mas agora estão extintas”. Indo um pouco além, os autores de [8] usam um exemplo interessante: “todas as espécies de plantas verdes terrestres que já viveram até o presente compartilham um ancestral comum: uma simples alga verde que viveu há 500 milhões de anos”. A pergunta chave é: Como foi possível atingir esse elevado nível de diversidade de espécies na história do planeta Terra, muitas vezes partindo de um único ancestral? Entender esses padrões de biodiversidade é o papel da Biogeografia, uma ciência multidisciplinar fundamentada em ecologia, biologia de

populações, sistemática, biologia evolutiva e ciências da Terra [8], [9] e [10]. Portanto, o que se busca com o estudo da Biogeografia é compreender e modelar processos espaço-temporais voltados para a geração e manutenção de diversidade. Espera-se que conceitos fundamentais da Biogeografia possam auxiliar a concepção de algoritmos que tratem de forma eficiente problemas de engenharia em geral, como otimização e mineração de dados.

Gerar diversidade, sob o ponto de vista de resolução de problemas é crucial em alguns casos. Como exemplos práticos, podemos considerar: a tomada de decisão em ambientes incertos [11], com a exigência de políticas alternativas a serem selecionadas em tempo real; otimização multimodal e multiobjetivo [12], com o requisito de escapar de ótimos locais de baixa qualidade e de povoar a fronteira de Pareto; de otimização dinâmica [13], com a necessidade de um comportamento exploratório capaz de promover uma rápida reação no caso de deslocamento da solução ótima; e uma ampla gama de problemas de aprendizado de máquina [14], caracterizados pela necessidade de alto desempenho e robustez em seus modelos matemáticos.

Obter soluções diversas não é uma tarefa trivial, sendo que técnicas distintas são empregadas nas mais variadas classes de algoritmos. A capacidade de lidar com múltiplas soluções de boa qualidade em paralelo é um dos diferenciais dos algoritmos bioinspirados, embora nem sempre a diversidade seja um mecanismo inerente às ferramentas bioinspiradas. Estas, muitas vezes, também não possuem um formalismo abrangente para sustentar os seus modelos computacionais. Estes são os principais diferenciais da Computação Biogeográfica: possui uma estrutura conceitual que descreve processos e definições e, por conseqüência, herda características presentes na Biogeografia, sendo a principal delas a evolução diferencial de espécies e indivíduos ao longo do tempo e do espaço.

Este artigo apresenta, de forma sucinta, a Computação Biogeográfica: explica brevemente os conceitos da Biogeografia, sua transição para a estrutura conceitual e, por fim, mostra um exemplo de algoritmo extraído de um subconjunto de processos.

## 2. Breves Conceitos da Biogeografia

No contexto da Computação Biogeográfica, o conceito de espécie é tratado de forma

primordial, vinculado ao conceito de espécie biológica, que define uma espécie como sendo uma população de organismos que apresentam isolamento reprodutivo em relação a outras populações, representando uma linhagem evolutiva separada. Adicionalmente, tem-se o conceito de habitat, o qual engloba uma ou mais espécies, no sentido de definir locais e condições ambientais em que o estabelecimento de populações é viável.

A Biogeografia estuda os padrões de diversidade de espécies e como surge a biodiversidade. Uma única espécie pode explorar todo um ecossistema, transpondo barreiras ecológicas, evoluindo e dando origem a diversas espécies descendentes, que podem ter diferenças genéticas em relação à espécie ancestral. Ao expandir sua amplitude, uma espécie poderá ou não se adaptar a novos ambientes. Se a adaptação for necessária, vários são os processos evolutivos que permitem a evolução e adaptação. As relações entre adaptação, evolução e o meio em que as espécies vivem podem ser explicadas pela taxonomia dos processos biogeográficos: os processos geográficos e ecológicos exercem uma pressão sobre o ecossistema. Caso as espécies necessitem de readaptação ao meio, os processos evolutivos se encarregarão da evolução. Se esta for um sucesso, novas espécies surgem. Caso contrário, a extinção é inevitável. A diversificação das espécies ao longo do tempo para preencher uma ampla variedade de habitats é denominada de **Radiação Adaptativa**.

Os processos que regem essa dinâmica espaço-temporal fazem parte da taxonomia de processos da Biogeografia, definida como:

- Processos Geográficos e Processos Ecológicos: (1) Vicariância e Dispersão; (2) Mudanças Ambientais; (3) Competição Interespecífica; e (4) Competição Intraespecífica.
- Processos Microevolutivos: (1) Reprodução; (2) Mutação; (3) Seleção Natural; (4) Adaptação; (5) Fluxo Gênico; e (6) Efeito do Fundador.
- Processos Macroevolutivos: (1) Especiação Alopátrica e Peripátrica; (2) Especiação Simpátrica; (3) Especiação Parapátrica ou Competitiva; (4) Extinção; e (5) Seleção de Espécies.

A principal diferença entre os processos Macro e Microevolutivos é a escala. A microevolução lida com mudanças evolutivas dentro de populações, que ocorrem como resultado de nascimentos, mortes e movimentos diferenciais de indivíduos com certas

características herdáveis. Em contrapartida, a macroevolução lida com mudanças frequentemente causadas pela proliferação diferencial e extinção de espécies, de forma que tais mudanças alteram substancialmente a diversidade das linhagens. Para informações detalhadas a respeito dos processos e padrões biogeográficos, consulte [8], [9] e [10].

Um conceito interessante que complementa a noção de adaptação a habitats distintos é o de superfície adaptativa de Wright [15]. Esse conceito descreve o mapeamento genético dos indivíduos em uma função de aptidão (*fitness*). A combinação de genes resulta em organismos pouco adaptados que estariam povoando os vales da função, enquanto que os mais bem adaptados povoam os picos da função. Cada pico equivale a uma região de alta aptidão para um determinado habitat ocupado por uma espécie qualquer. O domínio que compreende a função de aptidão delimita as possibilidades de adaptação da biodiversidade dentro de um ecossistema arbitrário.

### 3. Introdução à Estrutura Conceitual da Computação Biogeográfica

Processos centrais da Biogeografia são descritos abstratamente e em alto nível para compor a Computação Biogeográfica. A estrutura conceitual é composta através da extração de processos, definições e padrões da Biogeografia e da transposição destes através de formalizações matemáticas e estruturais. Permite, portanto, simular computacionalmente esses processos e definições, além de observar o comportamento evolutivo das espécies, no espaço e no tempo. Por fim, pode-se aplicar subconjuntos de processos objetivando a construção de algoritmos e ferramentas computacionais para a solução de problemas em várias áreas do conhecimento. Alternativamente, a Computação Biogeográfica pode ter o propósito de apenas simular processos e extrair comportamentos e padrões emergentes. Dado esse cenário tem-se a seguinte definição:

*A Computação Biogeográfica (CB) é uma linha de pesquisa que tem o objetivo de abstrair processos presentes na Ciência da Biogeografia e formalizá-los através de uma estrutura conceitual capaz de permitir o desenvolvimento de ferramentas computacionais voltadas para a solução de problemas complexos, além de simulações de processos e padrões da Biogeografia.*

Os seguintes processos estão propostos dentro da Computação Biogeográfica:

(1) Reprodução; (2) Mutação; (3) Fluxo Gênico; (4) Seleção Natural; (5) Competição Inter-específica; (6) Competição Intraespecífica; (7) Seleção de Espécies; (8) Especiação Simpátrica; (9) Especiação Competitiva; (10) Especiação Alopátrica; e (11) Especiação Peripátrica. A formalização matemática e estrutural da CB e de todos os processos pode ser conferida no seguinte trabalho: *Computação Biogeográfica: Formalização e Proposição de uma Meta-Heurística Espaço-Temporal* [16].

### 4. Algoritmo de Radiação Adaptativa

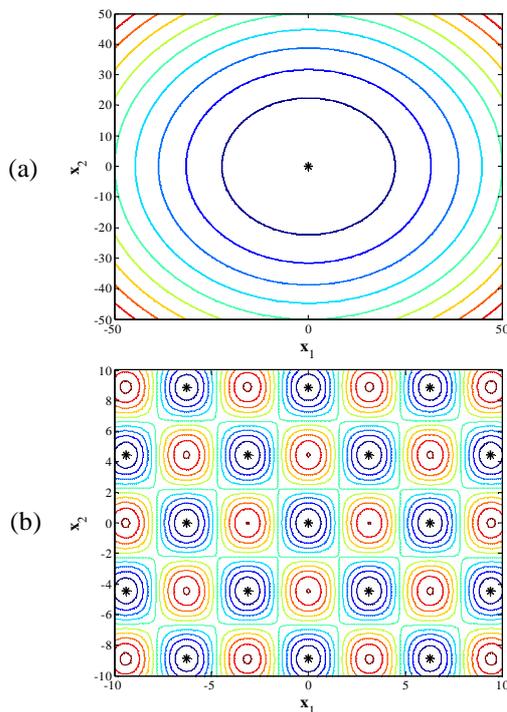
Com o objetivo de explorar a potencialidade da Computação Biogeográfica, a partir de um subconjunto dos processos descritos, um novo algoritmo foi proposto, denominado de *Adaptive Radiation Algorithm* (ARA). Como o próprio nome indica, ele simula os processos que regem a Radiação Adaptativa com objetivo de resolver problemas de otimização em espaços de busca contínuos ( $\mathfrak{R}^n$ ). Podem-se destacar características importantes do ARA:

- Geração e manutenção de diversidade de soluções;
- Ajuste automático do número de espécies e indivíduos;
- Retorno de um número de possíveis soluções de acordo com a demanda do problema;
- Ausência de parâmetros ajustáveis dependentes do problema;
- Auto-ajuste de processos internos; e
- Desempenho destacado em problemas multimodais;

O ARA opera sobre a superfície adaptativa correspondente à função que está sendo otimizada. Para isto utiliza os seguintes processos: Reprodução, Mutação, Fluxo Gênico, Especiação Simpátrica, Competição Inter-específica e Competição Intraespecífica. Uma primeira versão deste algoritmo pode ser conferida em [16].

Atualmente, o ARA encontra-se em uma versão mais robusta, composta de heurísticas e métodos determinísticos que controlam o surgimento de espécies de acordo com a superfície do problema. Adicionalmente aos processos da CB, o ARA possui duas buscas locais para refinamento de soluções. Os resultados preliminares mostram que o algoritmo é capaz de encontrar e manter uma única espécie por habitat (ótimo local) de acordo com a demanda do problema. Este é comportamento típico do ARA e é o que se espera em todas as suas execuções. Exemplos estão ilustrados nas

Figuras 1(a) e 1(b). A função convexa Sphere é otimizada com apenas uma espécie ao longo da execução (Figura 1(a)). O melhor indivíduo de cada espécie é representado por um “\*”. Em contrapartida, a função de Griewank (Figura 2(b)), altamente multimodal (a quantidade de ótimos locais é dependente do domínio e número de dimensões), é povoada por uma espécie em cada bacia de atração, representadas por habitats distintos. O ARA pode ser visto como uma representação computacional de um processo adaptativo que opera sobre a superfície adaptativa de Wright.



**Figura 1. Exemplos de execuções do ARA. (a) Função Sphere. (b) Função de Griewank. Cada “\*” representa o melhor indivíduo de uma espécie.**

As perspectivas futuras para o ARA são promissoras. Através da simulação da superfície adaptativa, dentro de um domínio pré-definido, é possível encontrar ótimos locais de acordo com a demanda do problema. Maximizando a exploração do espaço de busca, evitando redundâncias e economizando recursos computacionais. Essas características mostram que a exploração das particularidades que a Biogeografia possui pode trazer benefícios e diferenciais na concepção de algoritmos.

## Referências

- [1] L.N. de Castro, *Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications*, CRC Press, 2006.
- [2] L.N. de Castro, F. J. Von Zuben, *Recent Developments in Biologically Inspired Computing*, Idea Group Publishing, 2004.
- [3] I.N. da Silva, D. H. Spatti, D. H., E. A. Flauzino, *Redes Neurais Artificiais para engenharia e ciências aplicadas*, Artliber Editora, 2010.
- [4] T. Bäck, D. B. Fogel, Z. Michalewicz, *Evolutionary Computation 1 Basic Algorithms and Operators*, Institute of Physics Publishing (IOP), 2000.
- [5] T. Bäck, D. B. Fogel, Z. Michalewicz, *Evolutionary Computation 2 Advanced Algorithms and Operators*, Institute of Physics Publishing (IOP), 2000.
- [6] L.N. de Castro, *Engenharia Imunológica: Desenvolvimento e Aplicação de Ferramentas Computacionais Inspiradas em Sistemas Imunológicos Artificiais*, Tese de Doutorado, DCA-FEEC/UNICAMP, Campinas/SP, Brasil.
- [7] J. Kennedy, R. Eberhart, Y. Shi, *Swarm Intelligence*, Morgan Kaufmann Publishers, 2000.
- [8] J. H. Brown, M. V. Lomolino, *Biogeography*, 3rd ed., Sinauer Associates, 2006.
- [9] A. A. Myers, P. S. Giller, *Analytical Biogeography*, Chapman & Hall, 1991.
- [10] R. Hengeveld, *Dynamic Biogeography*, Cambridge University Press, 1990.
- [11] E. Czogala, H. Zimmermann, Hans-Jurgen, *Decision making in uncertain environments*, *European Journal of Operational Research*, Elsevier, vol. 23(2), pages 202-212, February 1986.
- [12] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*, Wiley, 2001.
- [13] Y. Jin, J. Branke, *Evolutionary Optimization in Uncertain Environments – A Survey*. *IEEE Trans. on Evolutionary Computation*, vol. 9, no. 3, 2005.
- [14] R. Pasti, L.N. de Castro, *Bio-inspired and gradient-based algorithms to train MLPs: The influence of diversity*, *Information Sciences*, v. 179, p. 1441-1453, 2009.
- [15] Wright, S, *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*, *Proceedings of VI International Congress of Genetics*, pp. 356-366, 1932.
- [16] R. Pasti, L. N. de Castro, F. J. Von Zuben, *Computação Biogeográfica: Formalização e Proposição de uma Meta-Heurística Espaço-Temporal*. In: XVIII Congresso Brasileiro de Automática, 2010, Bonito. XVIII Congresso Brasileiro de Automática, 2010.