

# 1

## Introdução

### 1.1

#### Motivação

Os problemas de otimização numérica são uma importante área de pesquisa, tendo aplicações em problemas como otimização de plantas industriais, mistura de produtos, refino, otimização de carteiras, mineração de dados e diversos outros (Back97, Michalewicz94).

Os algoritmos evolutivos têm sido uma importante ferramenta para a solução destes tipos de problemas. Estes métodos de otimização apresentam um bom grau de paralelismo durante o processo de busca, são facilmente adaptáveis a diversos tipos de problemas e têm bom desempenho em problemas de otimização ruidosos, descontínuos, não-diferenciáveis ou multimodais (Back97).

Os algoritmos evolutivos podem ser divididos em várias classes. Novas classes e novas técnicas dentro de cada uma das classes têm sido constantemente desenvolvidas. Alguns tipos de algoritmos evolutivos são:

- Algoritmos Genéticos (Back97, Michalewicz94)
- Programação Genética (Koza92)
- Evolução Diferencial (Storn95)
- Algoritmos Culturais (Reynolds94, Reynolds04)
- Programação Evolutiva (Yao99)

Apesar de serem utilizados com sucesso em diversos problemas de otimização, os algoritmos evolutivos apresentam, em algumas ocasiões, características que podem prejudicar o seu desempenho. Como estes algoritmos necessitam avaliar diversas vezes se as soluções encontradas por ele são adequadas, problemas onde essa avaliação seja computacionalmente custosa podem tornar proibitivo o uso dos algoritmos evolutivos. Além disso, os algoritmos evolutivos podem ter dificuldades para otimizar funções quando o cromossomo apresenta genes com problemas de epistasia (quando a “qualidade” de um gene é dependente de um outro gene).

Neste sentido, os algoritmos genéticos com inspiração quântica (Narayanan96, Han02, Han04) representam um dos mais recentes avanços na

área de computação evolutiva. Estes algoritmos se baseiam em idéias inspiradas na física quântica, em particular no conceito de superposição de estados, apresentando melhor desempenho em diversos tipos de aplicações. Mais especificamente, o algoritmo evolutivo com inspiração quântica usando representação binária (AEIQ- $\mathcal{B}$ ) (Han02) foi usado com sucesso em problemas de otimização combinatorial, apresentando resultados superiores em relação aos algoritmos genéticos convencionais em termos de tempo de convergência (e, conseqüentemente, em termos do número de avaliações necessárias para se atingir bons resultados). O algoritmo genético com inspiração quântica para otimização do problema do caixeiro-viajante (Narayanan96) também apresenta bons resultados quando comparado com os algoritmos evolutivos convencionais.

No entanto, nenhum dos algoritmos evolutivos com inspiração quântica preencheu uma lacuna importante: a otimização de problemas numéricos. Apesar do AEIQ- $\mathcal{B}$  poder ser utilizado para otimização neste tipo de problemas e apresentar alguns resultados satisfatórios (Han04), em geral, o uso de genes com codificação real, aliado a operadores específicos para esse tipo de codificação, produz resultados superiores, mais consistentes de experimento para experimento e com maior precisão numérica (especialmente em domínios grandes, onde a codificação binária requer uma representação proibitivamente longa) (Michalewicz94).

Um exemplo de problema onde os algoritmos evolutivos tradicionais não apresentam bom desempenho é na otimização de pesos para redes neurais. Por necessitarem realizar, em muitos casos, um número elevado de avaliações da função objetivo, os algoritmos genéticos tradicionais têm, em geral, um desempenho inferior aos algoritmos de aprendizado tradicionais usados em redes neurais (Yao99a, Ilonen03). Além disso, na área de redes neurais, normalmente necessita-se fazer diversas seqüências de aprendizado, devido ao fato de que a topologia ideal da rede neural não é, geralmente, conhecida a priori (Haykin99).

Um outro ponto importante é que, com exceção dos algoritmos culturais (Reynolds94, Reynolds04), os algoritmos evolutivos não mantêm, normalmente, conhecimento normativo sobre o espaço de busca de forma não-pontual. Em outras palavras, os algoritmos evolutivos não armazenam informações sobre as regiões mais promissoras do espaço de busca, a não ser através dos próprios indivíduos que formam a população em uma determinada geração. Esta informação armazenada nos próprios indivíduos é pontual e não pode ser compartilhada diretamente entre os indivíduos. Ao contrário, nos algoritmos culturais, esta informação é guardada no *espaço de crenças* e é diretamente compartilhada entre os indivíduos. Além disso, a informação não é pontual, mas representada por intervalos dentro do espaço de buscas que, ao longo do processo evolutivo, irá indicar as regiões mais promissoras do espaço de buscas. Este tipo de informação pode, como no caso dos algoritmos

culturais, melhorar o desempenho do algoritmo evolutivo consideravelmente. Esta informação também pode ser usada como semente em problemas de otimização *online* (onde a função que se quer otimizar varia ao longo do tempo), ao invés de (ou além de) usar um conjunto de indivíduos para este fim.

## 1.2

### Objetivos

Deste modo, baseado na discussão anterior, o objetivo principal desta tese é propor um novo modelo de algoritmo evolutivo, inspirado em paradigmas da física quântica, com as seguintes características:

- *Rápida Convergência* – Como a avaliação das soluções potenciais é, em geral, a responsável pela maior parte do tempo computacional gasto com o processo de otimização, deseja-se que o novo modelo seja capaz de realizar a otimização com um número menor de chamadas à função de avaliação do problema em questão;
- *Baixa Complexidade Computacional* – Em muitos problemas, o número de variáveis de entrada da função que se deseja otimizar é grande. Neste sentido, deseja-se que o algoritmo de otimização seja robusto com relação à dimensionalidade do problema, ou seja, que o aumento do número de variáveis não produza um aumento exponencial no esforço computacional ao se aumentar o número de variáveis da função que está sendo otimizada;
- *Conhecimento Normativo* – É interessante que o algoritmo de otimização seja capaz de armazenar conhecimento com relação às regiões promissoras do espaço de busca. Este conhecimento pode ser usado durante o processo de otimização convencional e em problemas de otimização *online* (onde a função que está sendo otimizada varia continuamente ao longo do tempo sem, no entanto, sofrer mudanças bruscas);
- *Aprendizado Compartilhado* – Como uma consequência do conhecimento normativo, é interessante que este conhecimento possa ser compartilhado entre os indivíduos da população que está sendo evoluída.

## 1.3

### Contribuições

Em função dos objetivos apresentados na seção anterior, este trabalho se concentrou na definição de um novo modelo de algoritmo evolutivo com inspiração quântica que ofereça as seguintes características:

- Representação específica para problemas de otimização numérica (representação por números reais) sem restrições;

- Maior velocidade de convergência quando comparado com algoritmos genéticos convencionais;
- Menor complexidade computacional com relação ao número de variáveis (dimensões) do problema que se quer otimizar;
- Mecanismos de compartilhamento de conhecimento entre indivíduos da população;
- Armazenamento de conhecimento normativo.

Além disso, este trabalho apresenta outras contribuições:

- Desenvolvimento de um ambiente de testes para o AEIQ–R usando a ferramenta MATLAB 7.0 nos ambientes Windows XP e Linux;
- Análise comparativa do desempenho do AEIQ–R com outros algoritmos evolutivos tradicionalmente usados em problemas de otimização;
- Análise do desempenho do algoritmo com relação à dimensionalidade do problema que se quer otimizar;
- Desenvolvimento de um algoritmo de treinamento de redes neurais recorrentes usando o AEIQ–R para problemas de previsão de séries temporais e para problemas de controle;
- Desenvolvimento de um ambiente de testes para o algoritmo de treinamento de redes neurais recorrentes usando a ferramenta MATLAB 7.0 nos ambientes Windows XP e Linux;
- Análise comparativa do desempenho das redes neurais treinadas com o AEIQ–R com outros algoritmos de treinamento de redes neurais e de aprendizado por reforço.

## 1.4

### Descrição do Trabalho

Este trabalho apresenta uma nova representação para os algoritmos genéticos com inspiração quântica. Esta representação, baseada em números reais, é uma poderosa ferramenta para a otimização de problemas numéricos e se mostra mais eficiente do que os algoritmos genéticos convencionais que usam o mesmo tipo de representação para otimizar funções matemáticas. Além disso, o algoritmo também se mostra mais eficiente na otimização de problemas numéricos do que o algoritmo evolutivo com inspiração quântica que usa representação binária. Entre outras propriedades importantes deste modelo está a sua capacidade de convergir rapidamente para uma boa solução usando poucos indivíduos na sua população. Isto reduz drasticamente o número de avaliações necessárias para a otimização, o

que é um importante fator de desempenho quando o modelo está sendo usado em problemas onde a avaliação consome muito tempo de processamento.

Além de apresentar o novo modelo, este trabalho também descreve uma série de testes realizados: alguns testes de otimização utilizando funções *benchmark* para otimização numérica, onde se pretende comparar o desempenho do algoritmo proposto com outros modelos de otimização inspirados na evolução natural, incluindo métodos já consagrados e métodos mais recentes; e alguns testes no treinamento de redes neurais (neuroevolução) recorrentes para previsão de séries temporais (aprendizado supervisionado) e para tarefas de controle (aprendizado por reforço).

Em especial, com relação ao uso do algoritmo para o treinamento de redes neurais, deseja-se mostrar que, ao contrário dos métodos de descida por gradiente, usados tradicionalmente para realizar o treinamento destas redes, e que se baseiam na otimização de uma única rede, o método de neuroevolução proposto é capaz de otimizar os pesos sinápticos de várias redes neurais ao mesmo tempo. Com isto, é possível diminuir a ocorrência de problemas de estagnação do processo de aprendizado por aprisionamento em mínimos locais, sem incorrer nos problemas tradicionais do uso de algoritmos evolutivos para otimização de pesos sinápticos já que, com o uso dos algoritmos evolutivos com inspiração quântica, o número de indivíduos que precisam ser avaliados para se atingir uma solução ótima é reduzido. Em outras palavras, com o uso de algoritmos evolutivos com inspiração quântica, pode-se acelerar o aprendizado ao mesmo tempo em que se evita uma convergência prematura para soluções sub-ótimas.

O método de neuroevolução proposto também apresenta outros benefícios que não estão diretamente ligados ao fato do algoritmo ter inspiração quântica. Estes benefícios se tornam evidentes, principalmente, quando se usa estes algoritmos para o treinamento de redes neurais totalmente recorrentes (*fully recurrent neural networks*), o que permite contornar problemas encontrados em métodos tradicionais de treinamento baseados em métodos de descida por gradiente para este tipo de redes tais como (Blanco01):

- Aprendizado Recorrente em Tempo Real (*Real Time Recurrent Learning - RTRL*) – este modelo tem como principal desvantagem o alto custo computacional para executar cada iteração;
- Retropropagação pelo Tempo (*Backpropagation Through Time - BTT*) – este modelo tem como principal desvantagem a necessidade de se saber o tamanho da seqüência de entrada à priori.

Finalmente, o modelo apresentado neste trabalho permite, entre outras coisas, a definição automática da topologia da rede neural final. Isto inclui, não só o número de processadores em cada camada, como também o número total de camadas que

compõem a rede neural. Além disso, devido à sua característica probabilística, o treinamento por algoritmos evolutivos com inspiração quântica evita a ocorrência de super-treinamento da rede neural, eliminando a necessidade do uso de conjuntos de validação para o processo de aprendizado.

Este trabalho tem como metas:

- Demonstrar que o novo modelo proposto tem um desempenho melhor do que os algoritmos evolutivos clássicos em diversos problemas de otimização numérica;
- Demonstrar a aplicabilidade e as vantagens deste novo modelo no aprendizado supervisionado de redes neurais em problemas de previsão e controle;
- Apresentar discussões sobre as características principais deste novo modelo e os resultados obtidos nos estudos de caso.

## 1.5

### Organização do Trabalho

Este trabalho está dividido da seguinte forma:

- No capítulo 2 é apresentado um resumo dos fundamentos teóricos necessários para a compreensão deste trabalho;
- No capítulo 3 o modelo proposto de algoritmo evolutivo com inspiração quântica e representação real é apresentado em detalhes;
- O capítulo 4 mostra os resultados obtidos em diversos estudos de caso;
- O capítulo 5 apresenta discussões sobre o modelo e os resultados encontrados, e aponta direções para trabalhos futuros.