

Trabalho Final de Computação Evolucionária

Uma exploração de planejamentos de percursos curriculares no escopo de algoritmos genéticos

Lorenzo Bicalho dos Santos Lopes
Engenharia de Sistemas
Universidade Federal de Minas Gerais
2020088155

Daniel Alves Resende
Engenharia de Sistemas
Universidade Federal de Minas Gerais
2020026834

Abstract—Esse documento apresenta o processo de estudo, desenvolvimento e aplicação do trabalho final da disciplina de Computação Evolucionária pela Universidade Federal de Minas Gerais no primeiro período de 2024. Este trabalho em particular propõe o resolvimento do problema de montagem de percurso curricular da grade atualizada do curso de Engenharia de Sistemas da UFMG. O problema foi aproximado usando algoritmos genéticos convencionais, com ênfase em sua modelagem. O código do trabalho foi armazenado em um repositório do GitHub para conveniência de acesso, acessível em: <https://github.com/dBorea/Comp-Evo-trabalho-final>

INTRODUÇÃO

A montagem de um percurso curricular é uma etapa fundamental e apresenta grandes desafios para estudantes de engenharia da UFMG durante sua jornada acadêmica. Segundo o relatório disponibilizado pelo setor de estatística da proreitoria de graduação [1], a taxa de eficiência do curso de Engenharia de Sistemas é de 6,6%. Ou seja, apenas uma pequena fração dos estudantes conseguem se formar dentro do período de integralização padrão do curso, em relação ao número de ingressantes do ano desses estudantes.

Portanto, este trabalho propõe um algoritmo que busca otimizar a construção do percurso curricular de um aluno de Engenharia de Sistemas, de forma a maximizar a taxa de aprovação em cada semestre. Foram estudados artigos que aproximaram o problema adjacente de *Timetable Scheduling* por meio de um algoritmo genético comum com ênfase na pré-modelagem heurística [1], um algoritmo que separa restrições entre "duras" e "desejáveis" para flexibilizar o problema [2], e um algoritmo evolucionário multiobjetivo [3]. Os algoritmos estudados serviram como base para a modelagem aplicada neste trabalho.

METODOLOGIA

Para a realização deste trabalho, foram consideradas as matérias obrigatórias da versão N-202402 do percurso curricular. Os códigos, nomes e pré-requisitos de cada matéria foram obtidos no relatório de percurso curricular disponibilizado pelo SIGAA. Os horários das matérias foram obtidos através do documento "Grade de horários para as atividades acadêmicas obrigatórias", disponibilizado no Moodle do colegiado de Engenharia de Sistemas.

As taxas de aprovação de cada matéria foram extraídas dos relatórios de avaliação do desempenho acadêmico do curso de

Engenharia de Sistemas [4], do ICEX [5], e da engenharia em geral [6], disponibilizados pelo setor de estatística da PROGRAD. Para as novas matérias da versão curricular, foi feita uma aproximação com base nas taxas de aprovação de matérias semelhantes.

O objetivo da implementação foi maximizar a média de aprovação de cada semestre. Para isso, foram consideradas as seguintes restrições:

- Não pode haver conflito de horários entre matérias de um mesmo semestre;
- Não podem ser cursadas matérias cujos pré-requisitos ainda não foram cumpridos;
- Não devem ser cursadas matérias de semestres demasiadamente distantes;
- Não pode-se deixar de cursar matérias com mais de três semestres de pendência.

Não foram consideradas matérias optativas, nem as matérias referentes a TCC1, TCC2 e estágio. Além disso, como no primeiro semestre o aluno não tem opção de alterar as matérias, as matérias do primeiro semestre não foram incluídas no algoritmo.

IMPLEMENTAÇÃO

A solução do problema foi a implementação de um algoritmo genético. Para cada semestre, o algoritmo é rodado quinhentas vezes e retorna as matérias recomendadas para aquele semestre. Essas matérias são consideradas como "concluídas" e não entram mais nas opções dos próximos semestres. O processo é repetido em loop até que todas as matérias estejam concluídas.

A. População Inicial

A população inicial é criada de forma aleatória. Cada indivíduo na população representa um conjunto de matérias a serem cursadas no semestre, codificadas como um vetor binário, onde 1 indica que a matéria está selecionada e 0 indica que não está. É gerado um vetor binário de tamanho igual ao número total de matérias disponíveis e criado uma população de tamanho igual a 100.

B. Fitness Function

A função de fitness avalia a qualidade de cada indivíduo com base em várias restrições e objetivos. A função é proje-

tada para maximizar a taxa de aprovação de cada semestre, minimizando conflitos e respeitando as dependências.

Primeiramente, cada matéria selecionada pelo indivíduo contribui positivamente para a fitness com base na taxa de aprovação, ponderada pelo número de horários alocados para aquela matéria. Esse valor é ajustado pelo quadrado da diferença entre o maior e o menor período das matérias selecionadas, penalizando trajetórias com grande dispersão de períodos e promovendo uma progressão mais linear ao longo do curso.

Em seguida, a função penaliza indivíduos que ultrapassam o limite máximo de horários permitidos por semestre. A penalização é proporcional ao excesso de horários, desencorajando a sobrecarga de matérias em um único semestre. Conflitos de horários entre matérias e a inclusão de matérias sem cumprir os pré-requisitos necessários também resultam em penalizações severas. Conflitos de horários são contabilizados ao verificar sobreposições entre os horários das matérias selecionadas, e dependências não atendidas são verificadas comparando os pré-requisitos das matérias com as matérias já concluídas pelo indivíduo.

Além disso, a função penaliza a diferença entre o semestre atual e o período mínimo de pendências, garantindo que matérias pendentes sejam priorizadas conforme o estudante progride em seu curso. Esse conjunto de critérios assegura que a função de fitness guie o algoritmo para encontrar soluções que sejam não apenas viáveis, mas também otimizadas para uma alta taxa de aprovação e uma distribuição equilibrada de carga horária e progressão curricular.

C. Seleção de Pais

A seleção de pais é realizada através do método de torneio. Nesse método, um conjunto de indivíduos é selecionado aleatoriamente e o melhor entre eles, com base na sua avaliação de fitness, é escolhido como pai. Esse processo é repetido até formar uma nova população. A seleção por torneio é vantajosa por ser simples de implementar e ajustar, permitindo a variação da pressão de seleção de acordo com o tamanho do torneio. Ela preserva a diversidade genética, pois indivíduos com menor fitness ainda têm chance de serem selecionados, o que é crucial para evitar a convergência prematura.

D. Crossover

O crossover combina dois indivíduos (pais) para gerar um novo indivíduo (filho). A função *crossover* realiza essa operação selecionando um ponto de corte aleatório e combinando as partes dos pais para criar o filho. Esse processo garante a troca de informações genéticas entre os pais, promovendo diversidade na população.

E. Mutação

A mutação introduz novas variabilidades na população, alterando aleatoriamente os genes de um indivíduo. A função *mutacao* percorre cada gene do indivíduo e, com uma probabilidade definida pela taxa de mutação, altera o valor do

gene. Esse processo ajuda a explorar novas soluções e evitar a convergência prematura.

F. Execução do Algoritmo

Para cada semestre, o algoritmo genético é executado quinhentas vezes. Em cada execução, os seguintes passos são realizados:

- 1) Inicialização: Uma população inicial de indivíduos é criada.
- 2) Avaliação: Cada indivíduo da população é avaliado usando a função de fitness.
- 3) Seleção: São selecionados novos indivíduos, a partir do torneio, como pais para a próxima geração.
- 4) Crossover: Novos indivíduos são gerados a partir dos pais selecionados.
- 5) Mutação: Os indivíduos gerados passam pela possibilidade de mutação para introduzir novas variações.
- 6) Substituição: A nova geração de indivíduos substitui a antiga.
- 7) Iteração: O processo é repetido por um número definido de gerações ou até que uma solução satisfatória seja encontrada.

Ao final de cada execução, as matérias selecionadas como concluídas são removidas das opções para os próximos semestres, garantindo que o progresso do aluno seja respeitado. O algoritmo continua até que todas as matérias obrigatórias tenham sido concluídas.

I. RESULTADOS

O algoritmo genético desenvolvido foi aplicado ao problema de construção do percurso curricular para um aluno de Engenharia de Sistemas da UFMG. O objetivo foi maximizar a taxa de aprovação em cada semestre, evitando conflitos de horários e respeitando os pré-requisitos das disciplinas. O período mostrado antes de cada matéria representa o semestre em que a matéria é apresentada no percurso curricular padrão. A seguir são apresentados os resultados obtidos para cada semestre:

1º Semestre

- Período 1 - MAT001: Cálculo Diferencial e Integral I
- Período 1 - QUI628: Química Geral E
- Período 1 - MAT038: Geometria Analítica e Álgebra Linear
- Período 1 - DCC203: Programação e Desenvolvimento de Software I
- Período 1 - ELE630: Introdução à Engenharia de Sistemas

Média de aprovação do semestre: 50.68%

Média de aprovação do semestre original: 50.68%

2º Semestre

- Período 2 - DCC217: Matemática Discreta para Engenharia
- Período 2 - ELT124: Sistemas Digitais
- Período 2 - FIS065: Fundamentos de Mecânica

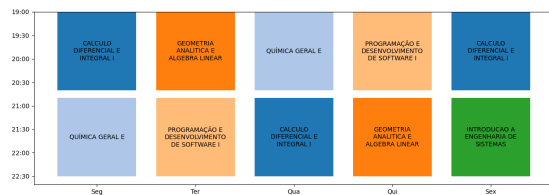


Fig. 1. Grade de Horários do 1º Semestre

- Período 4 - ELE632: Processos e Métodos em Engenharia de Sistemas

Média de aprovação do semestre: 59.53%

Média de aprovação do semestre original: 55.36%

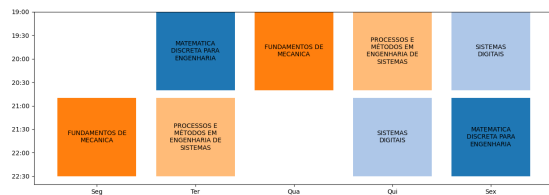


Fig. 2. Grade de Horários do 2º Semestre

3º Semestre

- Período 2 - DCC204: Programação e Desenvolvimento de Software II
- Período 2 - MAT039: Cálculo Diferencial e Integral II
- Período 3 - ELT029: Laboratório de Sistemas Digitais
- Período 3 - EST773: Fundamentos de Estatística e Ciência dos Dados

Média de aprovação do semestre: 65.42%

Média de aprovação do semestre original: 60.58%

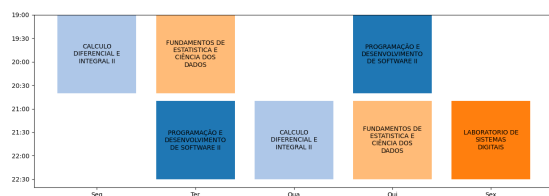


Fig. 3. Grade de Horários do 3º Semestre

4º Semestre

- Período 3 - DCC205: Estrutura de Dados
- Período 3 - ELE064: Análise de Circuitos Elétricos I
- Período 3 - FIS069: Fundamentos de Eletromagnetismo
- Período 4 - ELE631: Análise, Projeto e Programação Orientados a Objetos
- Período 4 - ESA019: Ciências do Ambiente
- Período 4 - MAT002: Cálculo Diferencial e Integral III

Média de aprovação do semestre: 55.47%

Média de aprovação do semestre original: 66.73%

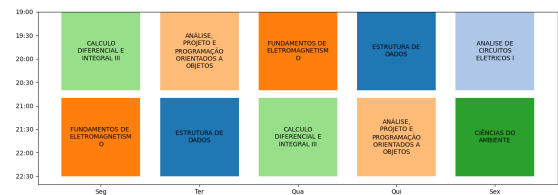


Fig. 4. Grade de Horários do 4º Semestre

5º Semestre

- Período 3 - MAT040: Equações Diferenciais C
- Período 4 - FIS086: Fundamentos de Oscilações, Ondas e Óptica
- Período 5 - EEE048: Fundamentos de Inteligência Artificial
- Período 5 - ELE028: Laboratório de Circuitos Elétricos I

Média de aprovação do semestre: 79.05%

Média de aprovação do semestre original: 80.26%

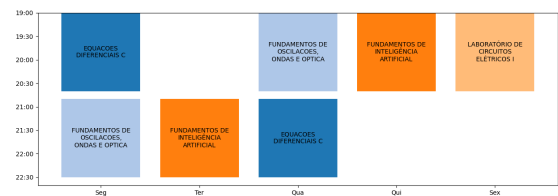


Fig. 5. Grade de Horários do 5º Semestre

6º Semestre

- Período 4 - ELE065: Análise de Circuitos Elétricos II
- Período 5 - ELT136: Fundamentos de Sistemas Dinâmicos e Controle
- Período 5 - EMT122: Introdução à Ciência dos Materiais
- Período 5 - ELE082: Pesquisa Operacional
- Período 6 - FIS152: Fundamentos de Mecânica dos Fluidos e Termodinâmica

Média de aprovação do semestre: 72.58%

Média de aprovação do semestre original: 66.06%



Fig. 6. Grade de Horários do 6º Semestre

7º Semestre

- Período 7 - EEE046: Sistemas a Eventos Discretos
- Período 7 - EEE049: Aprendizado de Máquina
- Período 7 - EEE050: Modelagem e Simulação Multifísica
- Período 7 - ELE088: Teoria da Decisão
- Período 7 - ELE633: Laboratório de Sistemas I

Média de aprovação do semestre: 77.62%

Média de aprovação do semestre original: 80.30%



Fig. 7. Grade de Horários do 7º Semestre

8º Semestre

- Período 6 - DCC011: Introdução a Banco de Dados
- Período 6 - ELE077: Otimização Não Linear
- Período 6 - ELT084: Dispositivos e Circuitos Eletrônicos Básicos
- Período 6 - ELT123: Arquitetura e Organização de Computadores
- Período 8 - ELE634: Laboratório de Sistemas II

Média de aprovação do semestre: 71.26%

Média de aprovação do semestre original: 86.27%



Fig. 8. Grade de Horários do 8º Semestre

9º Semestre

- Período 7 - ELT080: Laboratório de Circuitos Eletrônicos e Projetos
- Período 8 - DCC218: Introdução a Sistemas Computacionais
- Período 8 - EEE017: Confiabilidade de Sistemas
- Período 9 - EEE051: Laboratório de Gerenciamento de Sistemas

Média de aprovação do semestre: 88.83%

Média de aprovação do semestre original: 96.10%

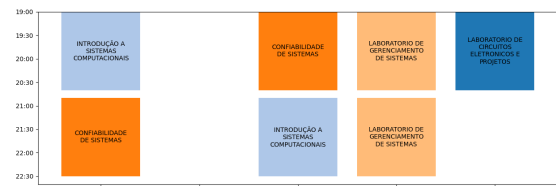


Fig. 9. Grade de Horários do 9º Semestre

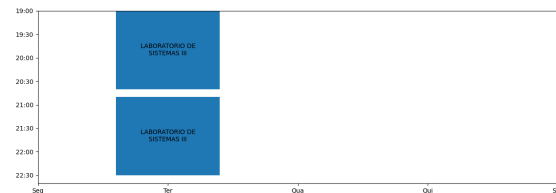


Fig. 10. Grade de Horários do 10º Semestre

10º Semestre

- Período 9 - ELE635: Laboratório de Sistemas III

Média de aprovação do semestre: 96.40%

Média de aprovação do semestre original: -%

A média geral de aprovação ao longo de todo o currículo foi de 71.68%. A média de aprovação do currículo original é de 71.37%.

DISCUSSÃO

Os resultados apresentados evidenciam que o algoritmo produziu uma distribuição com média semelhante à distribuição original. No entanto, ao examinar os horários de cada semestre, observa-se que a distribuição gerada pelo algoritmo favorece semestres com alguns horários livres. Isso ocorre devido à abordagem de maximização na função de aptidão, que considera tanto a taxa de aprovação quanto a quantidade de horários disponíveis para cada disciplina. Na prática, isso significa que, ao ter alguns horários livres, os estudantes têm mais tempo para estudar e realizar trabalhos de outras disciplinas, aumentando assim suas chances de aprovação.

Neste estudo, o algoritmo foi desenvolvido para otimizar os semestres individualmente. Por conseguinte, nota-se que nos primeiros semestres, a média de aprovação é superior à do percurso original. No entanto, devido à independência entre os semestres, à medida que o programa avança, as condições para alcançar uma otimização ideal começam a se tornar mais restritas, limitando, assim, o potencial da solução.

CONCLUSÃO

O processo delineado nesse documento evidencia alguns pontos chave da implementação de algoritmos genéticos para o resolvimento de problemas emergentes: a modelagem do problema e dos algoritmos responsáveis por seu resolvimento se mantém como um passo extremamente importante para

garantir o bom funcionamento das soluções apresentadas; algoritmos genéticos básicos apresentam diversas limitações que podem ser resolvidas com o aumento do escopo original do problema ou pela adoção de outros métodos de resolvimento por algoritmos evolucionários; resultados de sucesso parciais possuem valor analítico ainda que o objetivo final do escopo do projeto não tenha sido alcançado.

Os resultados apresentados nesse trabalho foram impactados por todos os pontos listados, e servem como um sumo exemplo de como a aproximação de um problema por meio de sua modelagem para um algoritmo genético pode ser complexa mesmo quando o problema se mostra caracteristicamente simples. Como fonte de aprendizado, este trabalho serve como um ótimo ponto de partida para os estudantes envolvidos, evidenciando os equívocos e sucessos do processo de desenvolvimento.

REFERÊNCIAS

- [1] SANTIAGO-MOZOS, R. et al. A two-phase heuristic evolutionary algorithm for personalizing course timetables: a case study in a Spanish university. *Computers Operations Research*, v. 32, n. 7, p. 1761–1776, jul. 2005.
- [2] HERATH, A. eGrove eGrove Electronic Theses and Dissertations Graduate School 2017 Genetic Algorithm For University Course Timetabling Problem. [s.l.: s.n.]. Disponível em: <https://egrove.olemiss.edu/cgi/viewcontent.cgi?article=1442context=etd>.
- [3] DATTA, D.; DEB, K.; FONSECA, C. Multi-Objective Evolutionary Algorithm for University Class Timetabling Problem. p. 197–236, 1 jan. 2007.
- [4] Oliveira, B. M., Teixeira, B. O. S., Pena, C. S. (2023). Avaliação do desempenho acadêmico e indicadores de evasão dos estudantes de graduação: Engenharia de Sistemas (Bacharelado). Belo Horizonte: Universidade Federal de Minas Gerais, Pró-Reitoria de Graduação/Setor de Estatística.
- [5] Teixeira, B. O. S., Flores, M. J. B. P., Gotelipe, L. G. de O., Silva, L. O. C., Vieira, L. P., Fonseca, B. B. (2022). Análise da retenção das atividades acadêmicas curriculares ofertadas e da evasão dos cursos sediados na Unidade Acadêmica: Instituto de Ciências Exatas. Pró-Reitoria de Graduação, Universidade Federal de Minas Gerais.
- [6] Oliveira, B. M. de, Teixeira, B. O. S., Gotelipe, L. G. de O. (2021). Análise da retenção das atividades acadêmicas curriculares ofertadas e da evasão dos cursos sediados na Unidade Acadêmica: Escola de Engenharia. Pró-Reitoria de Graduação/Setor de Estatística, Universidade Federal de Minas Gerais.