 <p>INSTITUTO FEDERAL ESPÍRITO SANTO Campus Serra</p>	<p>IFES</p> <p>Inteligência Artificial</p> <p>Exercício de Programação 2: Problemas de Otimização</p>	<p>Nota</p>
--	---	-------------

Professor: Sérgio Nery Simões

Data: 30/10/2023

Nome: _____

Turma: _____

Informações:

O trabalho deve ser submetido até a data limite na atividade do AVA. Deve ser submetido um arquivo compactado contendo (i) o código em Python dos programas desenvolvidos e (ii) o relatório em formato pdf.

É permitido que os alunos conversem a discutem suas soluções, mas os códigos-fonte devem ser produzidos individualmente. Se forem identificados casos de cola, serão abertos processos que podem culminar no desligamento do aluno.

Descrição do Trabalho

Contexto: Este projeto de programação consiste em comparar o desempenho dos algoritmos de otimização aplicados a diferentes problemas. Os resultados do projeto serão (i) o código desenvolvido para cada problema e (ii) um breve relatório apresentando e discutindo os resultados comparativos. Os algoritmos a serem comparados são:

- [HC-R] – *Hill-Climbing with Restart*;
- [SA] – *Simulated Annealing* e
- [GA] – *Genetic Algorithm*.

Desenvolvimento: Os algoritmos devem ser usados para tentar encontrar a solução ótima para problemas (conforme especificação). Na descrição do problema é definida a função objetivo e as funções para geração de vizinhos, e no caso específico do GA: mutação e *crossover*. Também são definidos os parâmetros que devem ser usados pelos métodos. Os algoritmos de otimização devem ser executados 30 vezes (no mínimo) e o relatório deve apresentar os resultados.

- (1) **TSP** (Travelling Salesman Problem – Problema do Cacheiro-Viajante);

Para ajudá-los a começar o desenvolvimento e prover um ambiente para testes dos algoritmos, foram fornecidos os seguintes códigos:

- (a) implementação do algoritmo **HC (clássico)** para busca local da solução aplicado ao problema TSP.
- (b) implementações parciais de **SA e GA** de busca global aplicado ao problema 8-Queens, com algumas funções para geração de gráficos comparativos.

A partir dos ambientes (a) e (b), os algoritmos deverão ser adaptados de acordo com as especificações de cada problema, fornecida mais a frente. Deve-se buscar implementar de forma estruturada, ou seja, desenvolver os algoritmos principais de forma mais macro e concentrar as alterações específicas de cada problema nas funções mais básicas.

Bom trabalho!

Problema

O problema de otimização consiste em encontrar uma solução para o **problema do caixeiro viajante** (*travelling salesman problem* - **TSP**). A entrada do TSP é um conjunto de pontos representando posições de cidades. A solução é o trajeto mais curto para, partindo de uma cidade inicial, visitar todas as cidades uma vez e depois retornar à cidade inicial. A Figura 3 exemplifica duas soluções para o TSP para um conjunto de pontos dado. Anexo à esta especificação são dados alguns conjuntos de pontos para realização de testes do algoritmo de otimização e um conjunto de pontos para o teste final e escrita do relatório. Os arquivos contêm as coordenadas (x, y) de um ponto por linha. Assuma que a cidade inicial do trajeto é dada pelo primeiro ponto do arquivo.

- **Função Objetivo:** A função objetivo a ser minimizada é o tamanho total do trajeto entre as cidades. A distancia entre duas cidades será obtida através da norma L2.
- **Representação do estado:** Uma solução para o problema será dada por uma sequência de números inteiros representando a ordem em que as cidades serão visitadas. Como cada cidade deve ser visitada apenas uma vez, não devem existir repetições de cidades no estado. A Figura 4 ilustra trajetos e o valor de estado de acordo com a representação proposta.
- **Geração de Vizinhos e Mutação:** A geração de vizinhos e a operação de mutação no algoritmo genético serão dadas pela troca da posição de duas cidades escolhidas aleatoriamente no trajeto. Por exemplo, o estado [2, 4, 5, 6, 3] poderia ser transformado em [2, 6, 5, 4, 3] trocando as posições das cidades 4 e 6.
- **Crossover:** Para evitar repetições de cidades após esta operação, usaremos um operador de *crossover* especial chamado **Order 1 Crossover** ou (OX Crossover). Explicações dele são apresentadas em [5] e [6].
 - **Justificativa:** a operação de *single-point crossover* **não** é adequada para estados representados como permutações de números porque ela pode gerar repetição ou ausência de números (cidades) nos estados filhos. Por exemplo, assuma que os estados pais são $p1 = [1, 3, 5, 2, 6, 4]$ e $p2 = [6, 4, 1, 3, 2, 5]$ e que a posição escolhida para *crossover* é a posição 1 (assumindo índices que começam de 0). Então, usando o *single-point crossover*, os filhos seriam $c1 = [1, 3, 1, 3, 2, 5]$ e $c2 = [6, 4, 5, 2, 6, 4]$. Observe que em $c1$, as cidades 1 e 3 são visitadas 2 vezes e as cidades 4 e 6 nunca são visitadas. Já em $c2$, as cidades 4 é visitada 2 vezes e a cidade 5 nunca é visitada.
- **Parâmetros:** O número de iterações assim como o tamanho população (10-30) e o número de gerações (ex: 50-300) do genético devem ser ajustados dependendo do número de cidades. Quanto mais cidades, mais difícil o problema e, portanto, mais chamadas à função objetivo podem ser necessárias.
 - **Importante:** para uma comparação justa, garanta que todos os algoritmos façam um mesmo número de chamadas à função objetivo. Teste diferentes número de restarts até encontrar um que leve a bons resultados no [HC-R] (*Hill-Climbing with Restart*). Para o [SA] (*Simulated Annealing*), defina a temperatura inicialmente como 1000 com fator de redução (ex: 0.99-0.9999) de forma a chegar próxima a 0 (zero) após 50-70% das iterações (ex: 2000-4000). Para o algoritmo genético, use probabilidade de mutação entre 10% e 20%.

Bom trabalho!

- **Visualização:** A visualização consiste em apresentar os pontos representando cidades e o trajeto dado pela solução atual. As Figuras 3 e 4 são exemplos desse tipo de visualização.

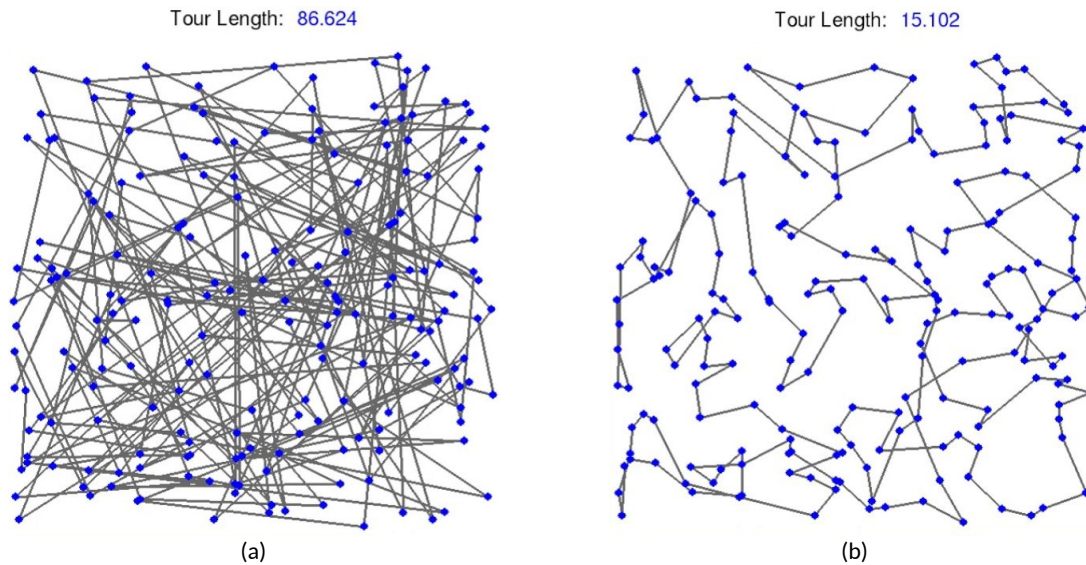


Figura 3: As figuras apresentam soluções para o problema do caixeiro viajante (*travelling salesman problem* - TSP). A solução (a) tem custo de caminho inferior à solução (b). Observe que o número de cruzamentos de trechos em (b) é menor que em (a).

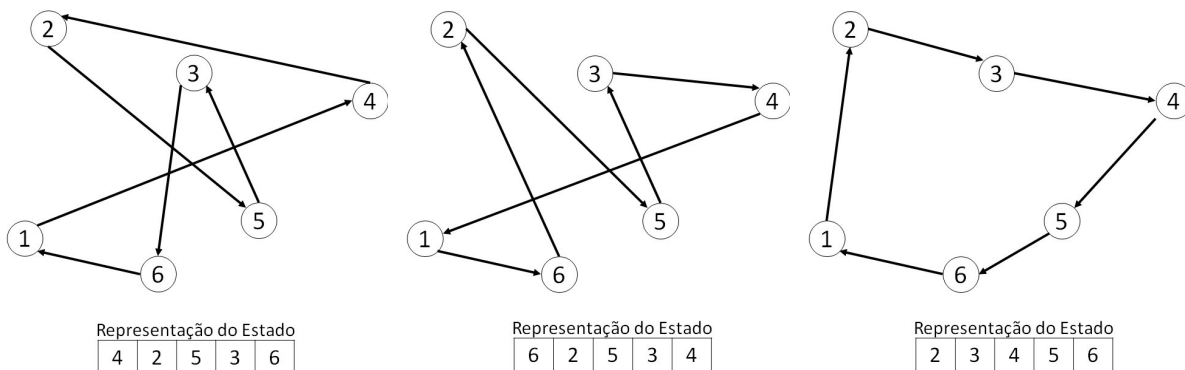


Figura 4: Exemplos de como as soluções do TSP serão representadas em nosso projeto.

Sugestões Gerais para a Realização dos Testes

- **Sugestão:** Não é obrigatório, mas bastante recomendável fazer um programa para rodar o processo de otimização. O programa receberia como entrada o nome do algoritmo de otimização e o nome do problema e geraria como saída um arquivo com os estados visitados e os valores da função objetivo (no caso do genético, os estados devem ser organizados em populações). Tendo estes arquivos, produzir o relatório será mais rápido e fácil. Sugestão: a biblioteca `argparse` [3] é uma boa alternativa para produzir interfaces de linha de comando.
- **Sugestão:** Não é obrigatório, mas recomendável fazer um *script* para fazer as 30 execuções programaticamente. Se vocês encontrarem bugs, podem ter que refazer os testes e ter que rodar tudo de novo manualmente toma bastante tempo.

Bom trabalho!

Importante mencionar que, para o desenvolvimento do trabalho, não é permitido o uso de bibliotecas que implementem os algoritmos, mas é permitido usar bibliotecas auxiliares, e.g., que implementem estruturas de dados. Esses algoritmos bem como sua aplicação em problemas de aplicações podem ser úteis em entrevistas de programação em empresas de grande porte, então considere este tempo de implementação como um investimento na sua carreira.

Formato do Relatório

O relatório deverá conter para cada problema, tabelas e gráficos resumindo os resultados obtidos e apresentando uma breve discussão.

- A discussão deve descrever os resultados (qual algoritmo obteve o melhor resultado, etc.) e prover uma explicação para o que foi observado com base nos conhecimentos adquiridos. Exemplo:

“O algoritmo de Hill-Climbing obteve desempenho inferior aos demais porque ficou preso em um mínimo local. Isso pode ser observado na Figura X que mostra a evolução da função objetivo. O valor da função objetivo parou de reduzir após N iterações”.

Os gráficos e tabelas a serem produzidos são os seguintes:

- Gráfico Box-plot com a sumarização dos resultados: Para comparar os resultados dos 3 algoritmos, plote um gráfico contendo os boxplots de cada um, ou seja, para visualização dos valores máximos, mínimos, percentis 25, 50 (mediana) e 75. Ver exemplo na Figura 1.
- Tabela com sumarização de resultados: Para cada algoritmo, separe o valor da função objetivo ao final de cada uma das execuções do processo de otimização. A tabela deve comparar os valores máximos, mínimos, medianos e o desvio padrão dos valores (ver ex. da Tabela 1).
- Gráfico da evolução da fitness em N iterações do algoritmo: Para cada algoritmo deve ser produzido um gráfico no qual o eixo X represente o número de chamadas à função objetivo e o eixo Y represente o melhor valor da função objetivo encontrado até aquele momento. A Figura 2 traz um exemplo do gráfico. Cada curva representa somente uma das execuções do algoritmo. Note que diferente do exemplo, a curva nunca subiria depois de descer, uma vez que estamos mostrando o melhor valor da função objetivo até aquele momento.

Pontuação: a pontuação será proporcional a realização das atividades propostas

<u>Atividades</u>	<u>Pontuação</u>
Implementação dos algoritmos	40 pontos [HC-R (10), SA (10) e GA (20)]
Relatório segundo especificação	60 pontos

Bom trabalho!

Exemplos de Figuras e Tabelas para serem usados na discussão dos resultados

Figura 1: Exemplo de gráfico com os boxplots dos custos finais após 10 execuções dos algoritmos A, B e C.

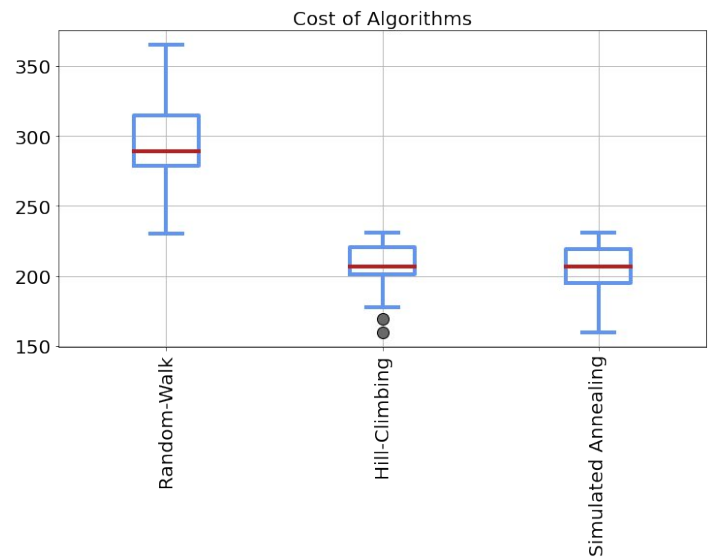
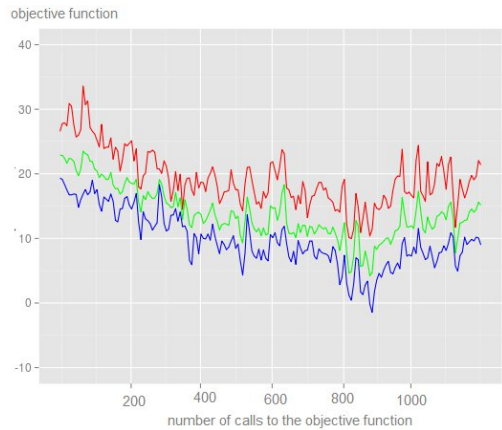


Tabela 1: Tabela de exemplo com estatísticas descritiva das execuções dos algoritmos de otimização A, B, C e D para o problema X.

Algoritmo	Max	Min	Mediana	Desvio Padrão
HC-R (<i>Hill-Climbing with Restart</i>)	3.442	0.123	1.235	1.982
SA (<i>Simulated Annealing</i>)	1.554	0.012	0.777	0.156
GA (Algoritmo Genético)	0.998	0.001	0.230	0.098

Figura 2: Exemplo de gráfico comparando a evolução da função objetivo nas 10 execuções de um algoritmo de otimização.



Referências

- [1] Use a função `numpy.random.randn` da biblioteca `numpy` para amostrar o valor aleatório.
<https://numpy.org/doc/stable/reference/random/generated/numpy.random.randn.html?highlight=randn#numpy.random.randn>
- [2] Use a função `numpy.random.uniform` da biblioteca `numpy` para amostrar o valor aleatório.
<https://numpy.org/doc/stable/reference/random/generated/numpy.random.uniform.html?highlight=uniform#numpy.random.uniform>
- [3] Biblioteca `argparse` para criação de interfaces de linha de comando.
<https://docs.python.org/3/library/argparse.html>
- [4] Use a função `seaborn.lineplot` (ou equivalente) para produzir o gráfico das curvas médias da função objetivo com os intervalos de variação.
https://seaborn.pydata.org/examples/errorband_lineplots.html
- [5] <https://www.youtube.com/watch?v=HATPHZ6P7c4>
- [6] Seção Order crossover (OX).
http://creationwiki.org/pt/Algoritmo_gen%C3%A9tico

Obs: Nos notebooks enviados também tem alguns exemplos de usos de algumas dessas funções para facilitar a solução dos problemas.