

## Introdução à Inteligência Artificial

Licenciatura em Engenharia Informática, Engenharia Informática – Pós Laboral e  
Engenharia Informática – Curso Europeu

2º Ano – 1º semestre

2020/2021

---

### Trabalho Prático nº 2 *Problema de Otimização*

#### 1. Objetivos

Conceber, implementar e testar métodos de optimização que encontrem soluções de boa qualidade para diferentes instâncias do problema a seguir descrito.

#### 2. Problema da Diversidade Máxima de Grupos

Neste problema pretende-se dividir um conjunto de  $M$  elementos em  $G$  subconjuntos. Cada um destes subconjuntos deverá ter o mesmo número de elementos ( $N$ ), i.e., deverá ter  $N = M/G$  elementos.

O objetivo da otimização é encontrar uma divisão que maximize a diversidade dos elementos pertencentes ao mesmo conjunto. A diversidade de um subconjunto  $G_1$  com  $N$  elementos é igual à soma das distâncias entre todos os pares de elementos que o constituem.

$$div(G_1) = \sum_{i=1}^{N-1} \sum_{j=i+1}^N dist(e_i, e_j)$$

Por exemplo, se o conjunto  $G_1$  contiver os elementos  $\{e_1, e_2, e_3\}$  então a sua diversidade é igual a:

$$div(G_1) = dist(e_1, e_2) + dist(e_1, e_3) + dist(e_2, e_3)$$

A qualidade de uma solução  $S$  é igual à soma das diversidades de todos os seus subconjuntos.

$$Qualidade(S) = \sum_{i=1}^G div(G_i)$$

O objetivo da otimização é encontrar a solução que **maximiza** este valor.

### Instâncias para teste

Estão disponíveis no moodle algumas instâncias deste problema para teste. O formato dos ficheiros txt onde está armazenada a informação de cada uma das instâncias é o seguinte:

- Linha 1: Dois inteiros indicando respetivamente o número de elementos (M) e o número de subconjuntos (G) a considerar. O resto da informação da linha 1 pode ser ignorado;
- As linhas seguintes contêm os valores das distâncias entre todos os pares de elementos que fazem parte do problema

#### Exemplo:

```
10 2 ss 5 5 5 5
0 1 69
0 2 100
0 3 73
...
7 8 65
7 9 68
8 9 39
```

Este exemplo tem 10 elementos e 2 subconjuntos.

Distância entre os elementos 0 e 1:  $dist(0, 1) = 69$ .

$dist(0, 2) = 100$

...

$dist(8, 9) = 39$ .

#### Exemplo:

Imagine que tem  $M = 6$  e  $G = 2$  com as distâncias definidas no ficheiro de teste fornecido (teste.txt).

À direita mostram-se duas soluções possíveis para o problema, S1 e S2, onde se dividiram aleatoriamente os **6** (M) elementos em **2** (G) subconjuntos: no subconjunto 1 (bit 0) e subconjunto 2 (bit 1).

Não esqueça que cada subconjunto deverá ter o mesmo número N de elementos (neste caso  $N = M/G = 3$ )

```
6 2 ss 1 1 1 1
0 1 10
0 2 12
0 3 5
0 4 10
0 5 2
1 2 10
1 3 15
1 4 10
1 5 5
2 3 5
2 4 4
2 5 10
3 4 5
3 5 20
4 5 10
```

S1 = [1 1 1 0 0 0]

S2 = [0 1 0 1 0 1]

A qualidade das soluções é calculada somando as diversidades do subconjunto 1 e do subconjunto 2:

**Para a Solução S1:**

$$\begin{aligned}\text{div}(G_1) &= \text{dist}(3, 4) + \text{dist}(3, 5) + \text{dist}(4, 5) = 5 + 20 + 10 = 35 \\ \text{div}(G_2) &= \text{dist}(0, 1) + \text{dist}(0, 2) + \text{dist}(1, 2) = 10 + 12 + 10 = 32 \\ \text{Qualidade}(S1) &= 35 + 32 = 67\end{aligned}$$

**Para a Solução S2:**

$$\begin{aligned}\text{div}(G_1) &= \text{dist}(0, 2) + \text{dist}(0, 4) + \text{dist}(2, 4) = 12 + 10 + 4 = 26 \\ \text{div}(G_2) &= \text{dist}(1, 3) + \text{dist}(1, 5) + \text{dist}(3, 5) = 15 + 5 + 10 = 30 \\ \text{Qualidade}(S2) &= 26 + 30 = 56\end{aligned}$$

Como se trata de um problema de maximização, a solução **S1** será melhor do que a **S2**

### 3. Métodos de Otimização

Neste trabalho pretende-se que implemente e avalie a capacidade de diferentes algoritmos de otimização para encontrar soluções de boa qualidade para o problema descrito. Sendo assim, no seu estudo deverá implementar os 3 métodos seguintes e efetuar um estudo comparativo aprofundado sobre o desempenho da otimização. No Moodle estão disponibilizados vários ficheiros de instâncias, com complexidades diferentes, para efetuar o estudo experimental dos vários algoritmos:

1. Algoritmo de pesquisa local (trepa-colinas, recristalização simulada ou outro, desde que aprovado pelo professor que lhe vai fazer a defesa);
  - Proponha uma forma de lidar com soluções inválidas: evitando que surjam na população, ou se surgirem, comparar estratégias de penalização ou de reparação;
  - Devem ser exploradas pelo menos duas vizinhanças diferentes.
2. Algoritmo evolutivo;
  - Proponha uma forma de lidar com soluções inválidas: evitando que surjam na população, ou se surgirem, comparar estratégias de penalização ou de reparação;
  - Devem ser explorados diferentes operadores genéticos.
3. Método híbrido combinando as duas abordagens anteriores.
  - Devem ser exploradas pelo menos duas abordagens híbridas diferentes usando os dois algoritmos implementados em 1. e 2.

### 4. Critérios de avaliação

- Originalidade e correção dos algoritmos implementados (30%)
- Experimentação e análise (50%):
  - Análise do impacto das componentes dos algoritmos: operadores, vizinhanças, arquiteturas híbridas, entre outros (20%).
  - Análise do efeito da variação de parâmetros (20%).
  - Outros testes (10%).
- Documentação e defesa (20%)

## 5. Normas de realização do trabalho prático

- O trabalho deve ser realizado **em grupos de dois alunos**.
- Os trabalhos serão sujeitos a defesa obrigatória, em data a anunciar.
- Cotação do trabalho: **6 valores**

## 6. Entrega do Trabalho

- Data limite para entrega: **22h00 do dia 17 de janeiro de 2021**
- Será dada uma penalização de **25%** por cada dia de atraso;
- Deverá ser submetido um ficheiro compactado, devidamente identificado com os nomes e números dos alunos que realizaram o trabalho. Esse ficheiro deverá conter o seguinte:
  - o **Relatório** abordando, pelo menos, os seguintes pontos:
    - Representação usada para o problema, descrição da função de avaliação e objetivo da otimização;
    - Descrição dos algoritmos e/ou das heurísticas utilizadas. Explicar quais as vizinhanças, métodos de seleção e operadores genéticos implementados;
    - Justificação das principais opções tomadas;
    - Resultados dos testes efetuados e respetiva análise. Os resultados a mostrar no relatório devem ser apenas um resumo baseado em tabelas/gráficos apresentando médias de várias repetições e respetivas conclusões. O estudo completo deve ser anexado num ficheiro Excel.
  - o **Código fonte** do programa, executável e exemplos para teste;
  - o **Ficheiro Excel** com os testes realizados para cada algoritmo.
- O trabalho deve ser entregue através da plataforma Moodle até à data limite indicada.
- As defesas serão nas aulas práticas da semana 18 a 22 de Janeiro e serão agendadas com os professores das aulas práticas.