

**Inteligência Artificial e Sistemas Cognitivos**

Mestrado em Engenharia Informática e Multimédia

**Relatório**

*Projeto*

Rodrigo Dias – 45881

# Introdução

# Inteligência artificial e sistemas cognitivos

# Projeto - Objetivo 1

Esta primeira parte foca-se em estudar o comportamento de redes neuronais com o objetivo de conseguir detetar padrões e semelhanças num conjunto de amostras para ser capaz de prever qual o resultado em dados nunca introduzidos. Esta abordagem baseia-se no comportamento de neurónios biológicos e a ideia é simular o comportamento deles, já que é destes que o cérebro do ser humano é composto e permitem a que sejamos capazes de raciocinar e detetar e reconhecer padrões.

Para este objetivo 1, são usadas técnicas de aprendizagem artificial supervisionada. Isto que dizer que, ao treinar a rede, indica-se qual o resultado esperado de forma a informar a rede se a previsão foi correta ou incorreta, ajustando assim os seus parâmetros até convergir para uma solução aceitável. Numa abordagem não supervisionada já não seria preciso indicar qual o resultado suposto da previsão.

## Realização de um protótipo com base numa plataforma de desenvolvimento de redes neuronais artificiais

### Aprendizagem da função lógica XOR

A picture containing diagram

Description automatically generatedA picture containing text, clock

Description automatically generatedPara realizar a função lógica XOR, pensou-se em implementar uma rede neuronal constituída por 2 entradas e 1 saída. Isto é problemático porque tendo um só neurónio, não existe maneira de criar uma fronteira para distinguir as combinações do XOR pois como se pode ver na F, a expressão de saída do neurónio gera apenas uma reta. A F ilustra os valores possíveis do XOR consoante os inputs e observa-se que uma reta não consegue separar os valores lógicos de 1 (a verde) dos valores lógicos 0 (a vermelho).

Uma solução a este problema seria ter 2 retas de maneira que a área entre elas (AND) possua apenas um valor lógico e a área fora (OR) possua o outro. Para tal, é possível utilizar 2 neurónios na mesma camada escondida, entre as entradas e a saída.

Diagram

Description automatically generatedComeçou-se por definir a rede neural com 2 entradas, uma camada escondida com 2 neurónios e 1 neurónio de saída. A F mostra como ficam ligados os neurónios entre si e os pesos (não mostra funções de ativação).

É possível pôr mais neurónios na camada escondida e isso leva a melhor desempenho na aprendizagem, mas decidiu-se medir o desempenho com o mínimo necessário, por isso espera-se que sejam precisas mais iterações para atingir um mesmo valor de erro. Deve-se ter em conta que não se deve colocar muitos neurónios nas camadas escondidas devido ao fenómeno de *overfitting*. Este fenómeno deve-se ao facto de haver muitos pesos para atualizar, levando ao modelo poder ficar demasiado bem treinado, passando, não a aprender, mas sim a memorizar. Quando depois se introduz dados que a rede nunca viu, iria cometer muito mais erros. Neste caso, como o número de entradas é limitado a 4, não seria problema.

A saída dos neurónios é ligada depois a uma função de ativação, podendo limitar os valores de saída através de uma função, como uma *sigmoid*, *tanh*, *relu*, entre outras. Escolheu-se usar a função de ativação *tanh* pois esta limita os valores (calculados pelos respetivos pesos) entre -1 e 1 e decidiu-se começar por fazer os testes da rede com uma configuração bipolar (o valor 0 dos bits é representado por -1) e caso se use a *sigmoid* por exemplo (que limita entre 0 e 1), não era possível obter -1 na saída. Esta função de ativação é importante no neurónio de saída, mas é possível usar outra função nos neurónios da camada intermédia, seria apenas uma questão de ver qual delas é que resultaria em melhor desempenho. Neste caso usou-se a mesma função de ativação (*tanh*) em todos os neurónios.

Usou-se um optimizador SGD (*Stochastic Gradient Descent*) que atualiza os pesos com o algoritmo de *back-propagation*, isto é, para cada iteração, coloca-se um dos conjuntos de valores de entrada na entrada, observa-se o que está à saída (que inicialmente há de ser aleatório pois os pesos são iniciados aleatoriamente) e define-se um valor de erro (neste caso o *mean-squared-error*) que é calculado subtraindo o valor na saída com o valor esperado (*target*) ao quadrado. Com esta medida, é possível calcular o gradiente para cada peso e subtrai-se o valor do peso, pelo valor do gradiente que corresponde a esse peso. Como o gradiente indica a derivada para qual o valor do peso maximiza o valor da função, como se está a falar de uma função de erro, quer-se minimizá-la, daí subtrai-se o valor da derivada correspondente. Esta subtração é feita a multiplicar por um valor de taxa de aprendizagem que permite aproximar mais rapidamente ou lentamente o valor do erro ao mínimo mais próximo. Isto é importante pois se o valor da taxa for muito pequeno, mais iterações são precisas para atingir um mínimo e no caso de ser muito alto, o valor do erro em vez de se aproximar do mínimo, estaria a afastar-se. Por isso, procura-se um valor intermédio que seja adequado.

Para tal, começou-se por medir a rede neural com *momentum* de 0 e com configuração de amostras de treino com ordem fixa e mediu-se o desempenho da taxa de aprendizagem. Registou-se 10 vezes, para cada taxa de aprendizagem, o número de iterações até o valor de erro passar abaixo de 0,1. Testou-se até 1000 *epochs* (iterações) e caso o valor de erro não fosse atingido, regista-se NA.

#### Efeito da taxa de aprendizagem



FIGURA0

É possível observar que há vezes em que o modelo não conseguiu atingir o valor de erro que era pretendido. É possível observar que a taxa de aprendizagem que mais levou a este problema foi a de 0,05, indicando que esta taxa é muito pequena, atualizando os pesos muito lentamente. Os restantes NA indicam que o modelo pode ter ficado preso num mínimo local.

#### Efeito da introdução de um termo de momento

Fez-se a mesma coisa, mas colocou-se um termo de momento de 0,5 e outro de 1. Isto serve para ajudar a “dar um empurrão” ao valor do erro para dar hipótese de sair de um mínimo local.



FIGURA



FIGURA

Com estes resultados, é possível observar que com momento de 0,5, o modelo conseguiu atingir o valor de erro pretendido com menos iterações que com momento 0, indicando que ajuda a sair dos vários mínimos locais que possa ter encontrado, indicando também que a taxa de aprendizagem mais adequada é a intermédia, por volta de 0,5.

Com momento de 1, repara-se que voltam a ser registados muitos mais valores NA, indicando que o valor de momento é demasiado grande e que está a “dar um empurrão” com muita intensidade.

#### Efeito da apresentação das amostras de treino com ordem fixa ou aleatória

Voltou-se a colocar o valor de momento a 0.5, pois foi o melhor dos testes previamente feitos, para agora testar o desempenho ao tornar a configuração com amostras de treino de ordem aleatória. São esperados ligeiramente melhores resultados pois ao treinar o modelo com ordem aleatória, exclui-se a hipótese de ele aprender qualquer padrão da ordem a que lhe são apresentados os valores de treino.



FIGURA

Pelos resultados obtidos, continua-se a ver que a taxa de aprendizagem perto de 0,5 é a mais adequada, pois tem uma média de iterações mais reduzida, e não tem valores NA comparando com os valores vizinhos. Em comparação com a F0, vê-se que a melhor taxa é a 0,5 porque possui um termo de momento, enquanto sem termo de momento, sem “empurrão”, a melhor acaba por ser taxa de 1, chegando ao mínimo mais rapidamente, mas sem maneira para sair de mínimos locais.

#### Efeito de uma codificação binária ou polar

Para este teste, voltou-se a pôr as amostras com ordem fixa e mudou-se para configuração binária. De maneira a comparar os resultados, não se alterou as funções de ativação, mantendo-as funções de *tanh*. No entanto faria sentido usar ativações como a *sigmoid* que limita os valores entre 0 e 1.



FIGURA

É possível observar que os resultados obtidos estão extremamente maus e isto deve-se exatamente ao estar a usar a ativação *tanh* permitindo uma maior gama de valores (entre -1 e 1) sendo mais complicado a rede neural entender que os valores devem ser ajustados entre 0 e 1.

### Aprendizagem de padrões de imagem

Com o objetivo de concretizar uma rede neural para deteção de padrões, implementou-se um modelo com 16 entradas (para os inputs dos pixéis de padrões 4x4) e 4 saídas, já que se terá 4 padrões diferentes para detetar, ilustrados na F. Caso seja detetado o padrão A, então as saídas estarão a [1,0,0,0], caso seja detetado o B, as saídas serão [0,1,0,0] e caso não seja nenhum padrão as saídas serão [0,0,0,0].



Como este é um problema complexo, criou-se uma rede neural com 32 neurónios na única camada intermédia e geraram-se amostras de treino com várias amostras A, B, C e D. Fez-se um algoritmo para gerar mais alguns padrões aleatórios, atribuindo o conjunto de saídas correspondente, tendo em conta que, apesar de gerado aleatoriamente, pode ter resultado num dos padrões definidos.

Chart

Description automatically generatedDesta vez utilizou-se um optimizador Adam com as configurações predefinidas pois resultou em melhor desempenho do que com aquele analisado no exercício anterior. Para treinar o modelo, considerou-se usar *validation\_split* para ajudar no desempenho, já que se carregam bastantes amostras de treino.

As funções de ativação aqui usadas foram a *sigmoid* à saída (pois os valores de saída são 0 ou 1) e a *relu* para a camada intermédia, pois foi a que deu melhores resultados após alguns testes.

Text

Description automatically generatedO processo é semelhante ao exercício anterior e após treinar o modelo com 200 iterações registou-se o desempenho do erro e da precisão para cada iteração, podendo observar na F que o erro está bastante baixo em contraste com a precisão, que está acima de 90%, indicando um desempenho muito bom na aprendizagem. Ao carregar depois o modelo, introduz-se um conjunto de padrões para ver o desempenho final e verifica-se que o modelo acerta nestes, como mostra a F.

## Aplicação de redes neuronais artificiais para resolução de um problema de escolha livre

A picture containing text, crossword puzzle

Description automatically generatedDecidiu-se fazer uma rede neural para resolver *Nonograms*, *puzzles* de lógica que consistem em preencher as células de uma matriz consoante os números indicados à esquerda das linhas e no topo das colunas, como ilustrado na F. No caso de uma linha ter os números “2 2”, indica que vão ser preenchidos dois pares de células da matriz separados por, pelo menos, uma célula em branco.

Para tal, o objetivo é colocar nas entradas da rede neural os números das linhas e colunas da matriz de forma a colocar na saída a matriz resolvida. Assumindo que se usará matrizes de 4x4, serão precisos 16 neurónios de saída. Para os neurónios de entrada, decidiu-se usar um para cada linha/coluna, caso contrário, o número de entradas vai depender do número de células preenchidas em cada linha/coluna. Imaginando que uma linha tem os números “2 1”, coloca-se no neurónio de entrada o valor 21, se tiver “1 1”, coloca-se o valor 11, se tiver “1”, coloca-se 1, etc. Dessa maneira, o número de neurónios de entrada serão 8, 4 para os números correspondentes às linhas e os outros 4 para os números das colunas.

Para treinar a rede neural, gerou-se primeiro um *array* de *arrays* únicos com 16 posições preenchidas com 1’s e 0’s aleatoriamente. Estes *arrays* de 16 posições são as matrizes quadradas transformadas de 4x4 para 1x16, que são já os valores que se espera ter à saída. Por isso, foi necessário um algoritmo que interpretasse essas matrizes para obter os números a serem introduzidos na entrada.

Após gerar as amostras configurou-se uma rede neural semelhante ao exercício anterior, mas com 3 camadas intermédias, cada uma com 1024 neurónios. A razão para esta configuração deve-se ao facto deste problema ser muito mais complexo e precisar de muitos mais pesos de forma a evitar *underfitting*, que é semelhante ao que acontece com o problema do XOR ao não introduzir pelo menos 2 neurónios na camada intermédia. Experimentou-se com valores maiores, mas tornava a complexidade muito maior e demorava muito mais tempo para treinar o modelo, sem um aumento muito significativo nos resultados. As funções de ativação foram semelhantes, na saída uma *sigmoid* e nas camadas intermédias *relu*, pois após alguns testes, esta foi a combinação que pareceu dar melhores resultados.

Após treinar o modelo, registaram-se os valores de erro e de precisão, tanto do conjunto de treino como o conjunto de validação (que contém 20% das amostras de treino) e, como mostra a F, observa-se que o modelo treinou até uma precisão acima de 90% e a Chart

Description automatically generatedprecisão do conjunto de validação perto dos 70%. Estes valores são distintos pois o conjunto de validação serve para ir testando o modelo enquanto treina e essas amostras são desconhecidas, logo tende a errar mais facilmente. A mesma coisa acontece para os erros.

Text

Description automatically generatedDe seguida testou-se o modelo com várias amostras e obteve-se valores próximos de 10% de erro e 80% de precisão, como se vê na F. Introduziu-se uma amostra *sample* para visualizar o resultado de saída e também é possível observar que o resultado foi o igual esperado.

# Projeto - Objetivo 2

Nesta segunda parte implementam-se algoritmos de raciocínio automático para otimização para resolver problemas como o N-Rainhas e o Caixeiro viajante. Este tipo de algoritmos tem como objetivo encontrar a (ou as) solução ótima a partir de um critério de avaliação ou função de desempenho para um determinado problema, de maneira a otimizar o valor obtido pela referida função de desempenho.

O código foi estruturado tendo uma classe para cada problema com os métodos adequados, uma classe para agrupar os algoritmos de *hill-climbing* estocástico, *hill-climbing* com reinício aleatório e *simulated annealing* e outra classe para agrupar as funções para o algoritmo genético, como ilustra a F.

Diagram

Description automatically generated

#### N-Rainhas

O problema das N-Rainhas consiste em encontrar uma configuração de um tabuleiro NxN com N rainhas dispostas nesse tabuleiro de maneira que nenhuma intercete com outra na horizontal, vertical ou diagonal.

O critério de avaliação para este problema será o número de colisões entre as rainhas, sendo que, quanto maior o valor, pior e pretende-se encontrar o estado com 0 colisões (solução ótima). Para evitar erros, transformou-se o número de colisões no seu simétrico de forma que quanto maior o valor, melhor a solução.

Cada estado é disposto num *array* de tamanho N em que cada *index* representa a posição da rainha no tabuleiro. Ou seja, cada *index* do *array* representa uma coluna do tabuleiro e o número nesse *index* indica a linha do tabuleiro em que a rainha está. Com este tipo de configuração diminui-se a complexidade forçando uma rainha por coluna, evitando só que haja colisões na horizontal e na diagonal.

Os estados vizinhos serão aqueles que, a partir de um estado atual, se move apenas uma das rainhas, ou seja, apenas se muda um dos números no *array*, movendo uma rainha apenas na vertical.

#### Caixeiro viajante

O problema do Caixeiro viajante consiste em encontrar o caminho entre as cidades que minimiza a distância percorrida sem percorrer duas vezes pela mesma cidade e voltando à cidade inicial.

Aqui o critério de avaliação é a distância total do percurso, sendo que, quanto maior o valor, pior e pretende-se encontrar um estado com a menor distância possível. Para evitar erros, calculou-se o simétrico para quanto maior o valor, melhor a solução.

Cada estado será composto por *tuplos*, que indicam as coordenadas de cada cidade, dispostos num *array*, sendo a ordem dessa disposição no *array* a ordem das cidades que serão percorridas.

Os estados vizinhos serão obtidos ao trocar uma cidade com outra qualquer. Ao calcular depois a distância, é colocado a cidade inicial no final do *array* calculando o percurso de volta à cidade que começou.

## Realização de uma biblioteca de métodos raciocínio automático para otimização

Neste ponto implementaram-se algoritmos que iniciam o problema a ser resolvido com uma configuração (estado) aleatória, a partir da qual se foi procurando a solução ótima, obtendo sempre um estado vizinho e avaliando se esse estado vizinho é melhor ou pior de acordo com o critério de avaliação do problema.

### Hill-Climbing estocástico

Neste algoritmo, a partir de um estado inicial aleatório, escolhe-se um estado vizinho que tenha a melhor avaliação de todos os vizinhos encontrados e repete-se até se encontrar um vizinho com estado pior ou até estar preso (encontrar os mesmos vizinhos várias vezes. Se isto acabar por acontecer é porque o estado chegou a um ótimo local e não consegue sair (preso) ou para sair só encontra estados com avaliação pior. Esta segunda opção é mais comum acontecer caso tenha encontrado a solução global, pois todos os outros estados são piores que o final, no entanto é possível encontrar um estado local cujos vizinhos sejam todos piores sem este ser a solução ótima.

### Hill-Climbing com reinício aleatório

Chart, surface chart

Description automatically generatedDe modo a resolver este problema de não encontrar solução ótima por ficar preso num ótimo local, implementou-se o algoritmo de *hill-climbing* com reinício aleatório. Este algoritmo basicamente faz o mesmo que o anterior, mas repete várias vezes e, para cada iteração, inicia com um estado aleatório. Isto faz com que se explore mais soluções a partir de pontos iniciais diferentes, dando a oportunidade a que algum deles encontre a solução ótima global. A F ilustra como este algoritmo ajuda a explorar os vários estados possíveis.

A cada iteração guarda-se o melhor estado encontrado, sendo mais provavelmente o estado ótimo, para no final apresentar como solução.

No entanto, se o problema for demasiado complexo há sempre hipótese de não chegar a encontrar um ótimo global. Outro problema é que para este tipo de algoritmo, é preciso procurar qual o estado vizinho que possui a melhor avaliação, levando a que se tenha de verificar todos os estados vizinhos possíveis e claro que, para problemas muito complexos, mais poder computacional será preciso e mais tempo irá demorar a encontrar qualquer solução.

#### A screenshot of a computer Description automatically generated with low confidenceN-Rainhas

Para resolver este problema, usou-se o algoritmo de *hill-climbing* com reinício aleatório com 25 iterações, com um tabuleiro com 8 rainhas. A F mostra o resultado final, verificando que não existem colisões entre rainhas.

#### Caixeiro viajante

Chart, scatter chart

Description automatically generatedChart, line chart

Description automatically generatedPara este problema, usou-se também o algoritmo de *hill-climbing* com reinício aleatório com 25 iterações, com 20 cidades num mundo de 100x100. A F mostra onde as cidades estão localizadas e a F mostra o percurso obtido e é possível observar que a ligação entre cidades está bem otimizada.

### Simulated Annealing

O algoritmo de *simulated annealing* foi pensado de forma a otimizar o desempenho dos algoritmos anteriores, não tendo de analisar todos os estados vizinhos possíveis para encontrar o melhor. Neste caso, o algoritmo escolhe um estado vizinho aleatório e caso o novo vizinho tenha melhor valor, atualiza o estado atual para continuar o processo. No entanto, caso seja pior, dá-se uma chance de atualizar o estado atual para esse estado pior. Isto deve-se ao facto de, caso o estado esteja preso num ótimo local, há possibilidade de andar para trás para o tentar evitar e possivelmente encontrar estados melhores.

Esta lógica é implementada dando a possibilidade a que no início se possa atualizar o estado atual por um pior, mas que ao longo do tempo esta hipótese diminua. Assim é possível explorar os estados iniciais mais aleatoriamente até ir reduzindo apenas aos melhores.

Esta implementação melhora drasticamente o desempenho pois procura um estado vizinho aleatório sem precisar de verificar todos os possíveis. No entanto, há possibilidade de não encontrar a solução ótima pois está a depender de uma fórmula de têmpera (*annealing*) e quando a temperatura chegar a 0, este processo de escolher o vizinho acaba, sendo o resultado o último estado atualizado. Este algoritmo é útil pois há um *trade-off* entre tempo de processamento e a solução mais ótima, ou seja, o tempo de processamento é reduzido drasticamente, mas a solução encontrada não é a melhor, no entanto, é bastante boa.

A fórmula de têmpera refere-se a que o algoritmo começa flexível (quente), permitindo escolher piores estados vizinhos, e ao longo do tempo torna-se mais restrito (arrefece), diminuindo a hipótese de os atualizar. Esta fórmula depende dessa temperatura e da diferença de avaliações entre o estado atual e o vizinho escolhido aleatoriamente. A temperatura vai diminuindo ao longo do tempo, diminuindo a probabilidade de escolher um vizinho mau, e a diferença de avaliações, quanto mais negativa for (se for positiva esta fórmula não é usada e o estado atual é logo atualizado), menor é a hipótese de atualizar com esse estado.

Chart

Description automatically generatedA chave está em escolher uma função de temperatura que permite maior flexibilidade ao início e no fim dê mais tempo para aprofundar. A função ideal seria algo semelhante à ilustrada na FTEMP.

#### A picture containing chart Description automatically generatedN-Rainhas

Desta vez usou-se o algoritmo *simulated annealing* com uma função de temperatura , que tem o comportamento igual à da FTEMP. A F mostra o resultado obtido por este algoritmo e vê-se que foi possível encontrar a solução ótima.

#### Caixeiro viajante

Para este problema, usou-se a mesma função de temperatura (FTEMP), com as mesmas cidades usadas no *hill-climbing* e observa-se na F que a solução não é tão boa como a outra, mas a distância também não é muito pior e a olho vê-se que a solução é bastante adequada, visto que o poder computacional necessário foi muito mais reduzido.

Chart, line chart

Description automatically generated

## Realização de uma biblioteca de algoritmos genéticos

Para este exercício foi implementado um algoritmo genético e este baseia-se em princípios da natureza como na seleção, cruzamento e mutação de indivíduos de uma população. Isto significa que, para um conjunto de indivíduos (estados) faz-se seleção de 2, tendo em conta uma função de adaptação (que dá maior probabilidade a indivíduos mais adaptados de serem escolhidos), faz cruzamento entre eles, ou seja, trocam-se parâmetros do estado entre eles para criar um novo, e com uma probabilidade baixa, realiza-se mutação para dar oportunidade de haver diversidade (com o objetivo de evitar soluções locais). Após estas operações, o indivíduo resultante é adicionado a uma nova população e o processo repete-se até gerar uma nova população com o mesmo número de indivíduos que a inicial. Com a nova população gerada, repete-se todo o processo um número de vezes suficiente até obter uma solução adequada.

Este algoritmo é ótimo pois caso haja uma alteração no mundo os indivíduos adaptam-se de uma maneira muito eficiente, ao contrário dos algoritmos apresentados anteriormente, que acabam por aprofundar num certo objetivo.

Para pôr o algoritmo a funcionar é necessário ter em conta o tipo de problema que se pretende resolver pois, por exemplo, no caixeiro-viajante não pode haver cidades repetidas, então tem de se implementar um algoritmo de mutação e de cruzamento que tenha isso em conta, ao contrário do N-Rainhas que pode trocar apenas um número (dentro do limite do tabuleiro).

Para tal, fizeram-se os métodos de reprodução (cruzamento), função de adaptação e a função para gerar a população inicial. Uma coisa que não foi feita e seria mais adequado fazer era também fazer a função de mutação como método de cada problema pois o tipo de mutação é diferente no exemplo dos problemas que se está a resolver, onde no N-Rainhas pode haver indivíduos com genes repetidos e no Caixeiro viajante não. Esta função de mutação foi feita trocando dois genes do indivíduo, o que é permitido em ambos os problemas abordados.

O método de seleção foi do tipo roleta, em que se escolhe um dos indivíduos dando maior probabilidade aos indivíduos com melhor valor de adaptação. Isto permite a que sejam mais facilmente escolhidos os melhores, dando oportunidade de escolher alguns piores para diversificar e evitar soluções locais.

Outro método feito foi o de melhor seleção que basicamente escolhe o melhor indivíduo dentro de uma população de acordo com a função de adaptação e será usado apenas no final do processo do algoritmo para apresentar a melhor solução obtida.

A função de adaptação foi feita com base no critério de avaliação já feito para os algoritmos anteriores, apenas adaptando a valor para ficar entre 0 e 1, sendo 1 o melhor possível. Para tal, usou-se a fórmula sendo o *value*, o valor absoluto do critério de avaliação (pois está negativo) e que quanto maior for (mais colisões ou maior distância), pior é o valor de adaptação. Para o caso do N-Rainhas é possível obter um valor de adaptação de 0, mas no Caixeiro viajante já não, pois só aconteceria caso a distância fosse 0. No entanto, basta que sejam diferentes para permitir maior probabilidade de seleção aos melhores indivíduos.

#### N-Rainhas

A screenshot of a computer

Description automatically generated with low confidenceA picture containing shape

Description automatically generatedNeste problema carregou-se uma população inicial com 100 indivíduos gerados aleatoriamente, cada um com uma configuração de 5 rainhas, e fez-se 100 gerações (iterações) com o algoritmo genético. A razão para se fazer com 5 rainhas e não com 8 como nos algoritmos anteriores, deve-se ao facto de este ser apenas pegar em indivíduos que foram gerados aleatoriamente e trocar informação entre eles, ou seja, este processo acaba por ser bastante aleatório. A única parte menos aleatória é a de seleção que tem em conta aqueles com melhor valor. Isso não indica que ao fazer cruzamento se encontre uma solução melhor, aliás até pouco provável. Por isso, para algoritmos com muita complexidade, seriam precisos maior número de gerações e de elementos na população, levando a muito mais tempo e poder computacional para obter um resultado.

A F mostra o resultado obtido para esta configuração e a F mostra para uma configuração mais complexa (com 8 rainhas, mantenho o tamanho da população e de gerações) e observa-se que conseguiu chegar perto, mas não à solução ótima.

#### Caixeiro viajante

Neste problema carregou-se uma população com 100 indivíduos também gerados aleatoriamente sobre as mesmas cidades, ou seja, cada indivíduo tem uma ordem aleatória do percurso, com 5 cidades num mundo de 100x100.

Chart, line chart

Description automatically generatedChart, scatter chart

Description automatically generatedA F mostra as cidades a serem visitadas e a F a ordem do percurso e vê-se facilmente que o resultado obtido foi a solução ótima.

Chart, scatter chart

Description automatically generatedChart, line chart

Description automatically generatedA F mostra as cidades para o problema mais complexo (com 10 cidades) e observa-se na F que a solução obtida não foi a melhor apesar de não muito má.

## Estudo e utilização de uma outra plataforma de algoritmos genéticos

Para esta parte utilizou-se a biblioteca de algoritmos genéticos *PyGAD* e após analisar a documentação e a implementar para os problemas a resolver, reparou-se que o algoritmo é muito mais flexível, isto é, é possível usar vários tipos de operações como cruzamento ou seleção, sem ter de implementar o código pois a biblioteca abstrai essa lógica, pedindo apenas as restrições que esses métodos devem ter, como a questão de não haver genes repetidos e o tipo de seleção ser por roleta ou uniforme, entre outras opções.

A biblioteca também permite introduzir um parâmetro de paragem que indica ao algoritmo para parar quando um indivíduo atingir um valor de adaptação específico ou quando saturar um número de vezes, isto é, não ter progresso por algumas gerações.

Reparou-se também que a única função que tem de ser introduzida é a função de adaptação, pois esta pode variar entre problemas e é aquela que mede o desempenho de cada indivíduo para os poder definir um melhor que outro.

Também permite gerar a população inicial dando informação da quantidade de genes, o tipo e quais os valores permitidos. Por outro lado, também é possível introduzir a população inicial que resulta no mesmo resultado, facilitando todo este processo.

Também permite vários tipos de crossover como *single point* ou uniforme. O tipo *single point* escolhe um ponto entre os genes de dois indivíduos e cria um novo juntando a primeira parte de um com a segunda do outro (o mesmo feito na biblioteca anterior) e o uniforme é a escolha de um dos genes dos indivíduos pais, para gerar o filho.

#### N-Rainhas

Para este problema, indicou-se o número de gerações como 100, 2 indivíduos para cruzar e usou-se a mesma função de adaptação que foi feita no exercício anterior, de modo a poder comprar com a biblioteca feita. Também se indicou a população inicial, os valores que os genes podem tomar, a probabilidade de mutação (também 10% como anteriormente), o método de seleção como roleta, cruzamento *single point* e o tipo de mutação aleatória, permitindo genes repetidos.

Reparou-se que algumas vezes a solução final não era a ótima, mas ao colocar o parâmetro de paragem do algoritmo ao atingir um valor de adaptação de 1 (máximo), já encontra sempre uma solução ideal. Assim que encontra um indivíduo com esse valor, interrompe o algoritmo, informando essa solução.

De maneira a comparar à biblioteca feita, removeu-se o parâmetro de paragem. Observou-se que muitas vezes não é encontrada uma solução, mas na biblioteca realizada, também demora bastante mais tempo até fazer 100 gerações. Isto indica que o algoritmo feito não está tão otimizado, mas se calhar é por isso que mais facilmente é encontrada uma solução.

#### Caixeiro viajante

Este problema não foi possível fazer para esta biblioteca pois as cidades são representadas por *tuples* de coordenadas (X, Y) e a biblioteca não permite isso. Tentou-se adaptar ao problema, mas não foi possível encontrar uma solução.

No entanto, os parâmetros que seriam necessários especificar eram os de que não podem existir genes repetidos (cidades) e o tipo de mutação é *swap*, a troca entre a ordem das cidades. O critério de paragem não pode ser até atingir um valor de adaptação pois não se sabe qual a menor distância possível de percorrer e por isso é mais correto definir a saturação, ou seja, caso o melhor indivíduo seja o mesmo por um número de gerações, acaba e assume esse como a solução final.

# Projeto - Objetivo 3

Esta parte aborda as técnicas de aprendizagem por reforço e raciocínio automático para planeamento.

A aprendizagem por reforço consiste num agente que interage num ambiente e verifica o reforço que obtém após executar uma ação num certo estado, isto é, após uma ação sobre um estado, o agente observa se obteve recompensa pela ação ou se foi castigado, tentando entender quais as melhores ações até atingir um estado final de uma maneira eficiente, ou seja, com a melhor recompensa possível.

O problema começa com o agente num estado inicial e, sem ter qualquer conhecimento do mundo, executa uma ação aleatória, observa a recompensa e guarda informação numa memória de maneira a “lembrar” se já executou uma ação num certo estado e qual era a melhor decisão a tomar. Inicialmente esta memória começa vazia e por isso o agente começa por executar ações aleatoriamente, sem saber para onde ir.

O processo de seleção da ação dá uma chance ao agente de ir explorando ou aproveitando, isto é, com uma probabilidade, o agente executa uma ação completamente aleatória ou uma ação que maximiza o valor da recompensa, segundo a memória que tem guardada. Este valor de probabilidade deve ser maior para a opção de aproveitamento (*greedy*) mas dando chance de também poder explorar. Isto permite o agente “aventurar-se” por estado que nunca conheceu.

Outro fator importante é a recompensa que deve ser atribuída a cada ação. Tem de haver uma maneira de dizer ao agente que quando está a bater num obstáculo, para não o voltar a fazer. Isto é feito com uma recompensa negativa sempre que ele toma uma ação num estado que o leve a atingir um obstáculo. Outro aspeto em ter em conta é o custo do movimento. Caso a recompensa por cada movimento normal seja nula, o agente pode acabar por ficar preso pois não sabe que está a fazer algo errado, apesar de também não estar a atingir o alvo necessário. Para tal, coloca-se uma recompensa muito pequena por cada movimento, desta forma, como inicialmente não tem muita informação na memória, quando decide tomar a opção *greedy*, ele vê que num certo estado já tomou uma ação, que lhe deu uma recompensa pequena, mas negativa motivando-o a tomar uma ação que ainda não foi feita.

Desta maneira, o agente explora o mundo até conseguir atingir o alvo, que lhe dá uma recompensa positiva (positiva o suficiente para indicar que vale a pena chegar ao objetivo para compensar todas as recompensas negativas do custo do movimento até chegar ao alvo). Quando atinge o alvo, o agente volta ao estado inicial para começar a aperfeiçoar o caminho mais eficiente, isto é, com base na memória que foi atualizando, vai ajustando sabendo quais as decisões tomar, tendo sempre hipótese de tomar uma ação aleatória para explorar um caminho possivelmente mais eficiente.

O raciocínio automático para planeamento tem como objetivo verificar qual a sequência de ações que levam o agente a chegar a um estado final também de uma maneira eficiente. A diferença é que o agente faz uma procura do alvo podendo fazer uma procura em árvore. Esta consiste em ir expandido e explorando nós (estados) com operadores (ações). Alguns dos métodos desta procura são a procura em profundidade, em largura e de custo uniforme.

O primeiro é uma procura que expande os nós assim que encontra um, apenas expandindo um nó adjacente caso não tenha encontrado a solução no que começou por expandir. Isto leva ao problema que, quando encontrar solução, pode não ser pelo caminho mais eficiente.

O segundo é uma procura que expande os nós adjacentes e só depois de explorar todos é que expande os nós da camada abaixo. Desta forma, quando for encontrada a solução, será pelo percurso ótimo, mas requer reter muito maior memória e tempo de processamento.

O terceiro é uma procura de custo uniforme que expande nós pela ordem de um custo que é atribuído ao aplicar o operador a um determinado estado. Assim seria possível explorar com prioridade os nós que mais eficientemente atingem o alvo, com menor custo possível.

A procura do tipo melhor-primeiro tem em conta este custo uniforme, mas aproveita também uma estimativa de custo, uma função heurística, do nó atual até ao alvo.

## Realização de um agente capaz de navegar num espaço de dimensões discretas, desviando-se dos obstáculos e recolhendo o alvo

### Raciocínio automático para planeamento, com base no método Wavefront

Este algoritmo tem como base a ideia de expansão de nós com um tipo de procura em largura. Isto significa que a partir do estado alvo procuram-se os estados adjacentes e atribui-se um valor mais pequeno quanto mais distante estiver do alvo. Percorre-se todo o mundo para obter todos estes valores do mundo. De seguida, a partir de qualquer estado, aplica-se um algoritmo semelhante ao *hill-climbing* tendo como objetivo procurar o melhor estado vizinho, ou seja, o estado vizinho com maior valor atribuído. No entanto, pode-se aproveitar o método de procura de estados adjacentes, escolhendo aquele com maior valor. Como isto faz uma procura por frente de onda, não é preciso preocupar com máximos locais pois o algoritmo permite ir diretamente ao alvo, a partir de qualquer ponto do mundo.

A enorme desvantagem deste problema é que tem de guardar uma memória para o mundo inteiro e tem de o conhecer e caso o mundo altere, digamos a alterar o alvo para outro quando o primeiro é atingido, todo o processo tem de ser feito outra vez para percorrer a frente de onda para o novo.

### Raciocínio automático para planeamento, com base no método de procura em espaços de estados RTAA\*

Neste algoritmo faz-se um tipo de procura informada, isto é, é conhecido o domínio do problema, tendo maneira de obter uma estimativa do quão longe está o alvo (função heurística) e, com o custo do percurso desde o nó inicial até ao final, calcula-se uma estimativa do custo do percurso total. Existe também uma procura de custo uniforme que tira partido apenas do custo até ao nó atual, como referido no início deste objetivo. Outra procura é a sôfrega que tem em conta apenas a estimativa do custo do nó atual até ao alvo. Para este exercício usa-se a combinação dos dois tipos.

Começa-se por um nó inicial e expandem-se os nós que minimizam este custo total tentando obter o caminho no qual levará mais eficientemente o agente até ao objetivo. Para isto seria necessária uma memória de nós explorados e de nós na fronteira, tendo cada nó informação do nó antecessor, do estado que representa e do operador que o antecessor efetuou. É também necessário dar informação do tipo de procura, que para este caso seria a procura A\*.

Para o modelo do problema é preciso indicar quais os operadores possíveis, um método para verificar quando o nó representa o alvo e a função de heurística para indicar a distância que um determinado está.

Os métodos deste problema serão utilizados no mecanismo de procura que determina a lógica para expandir os nós respetivos, enquanto existir fronteira e enquanto a profundidade máxima introduzida ainda não tiver sido atingida. Isto permite poupar memória indo procurando até uma certa profundidade limitada, voltando a correr o processo, desta vez a partir do último nó que melhor heurística possui. A desvantagem é que o percurso encontrado pode não ser o mais eficiente possível pois o agente pode estar a ir na direção de um beco sem saída e se a profundidade não permitir detetar isso, na iteração a seguir, ele irá voltar para trás.

Nota: Este algoritmo não chegou a ser implementado.

### Aprendizagem por reforço com base no algoritmo Dyna-Q

Este algoritmo é feito utilizando um mecanismo de aprendizagem por reforço que é inicializado com as opções de ação que podem ser tomadas (neste caso para movimentar o agente), uma memória esparsa para guardar pares estado-ação com o valor de ação (Q) de cada um, obtendo um valor de omissão caso um certo par ainda não exista, um método de seleção de ação (que aproveita (*greedy*) ou explora, como referido no início deste objetivo) que escolhe as ações com base no conteúdo na memória e um algoritmo de aprendizagem por reforço, sendo neste caso o *Dyna-Q*.

É aqui que são feitos os passos de aprender, atualizar a memória com o que aprendeu e de simular sobre um modelo do mundo, atualizando também a memória com os dados dessa simulação. O modelo do mundo guarda a informação de qual o estado que se segue ao atuar com uma ação específica e qual o reforço obtido.

Para a aprendizagem, é usada a memória que guarda o valor de ação (valor Q), representada na F. Este valor Q é atualizado baseado na ação que maximiza o valor Q da iteração seguinte, o reforço e o Q do próprio estado e isto serve para prever qual o resultado da ação que o agente irá tomar e se é melhor a longo prazo. Assim o agente pode ver que é melhor passar por uma zona com menos recompensa, mas que no final do problema este acaba por ser o percurso que mais compensa. É usado um valor de fator de aprendizagem (α) que indica o quanto deve ter em conta esta hipótese. Outro valor usado é o fator de desconto (γ) que desconta a previsão da ação que maximiza o valor Q do estado seguinte.

O fator de aprendizagem foi colocado a 0,7 e o fator de desconto a 0,95. O fator de desconto é alto pois quer-se aproveitar a previsão do estado seguinte. O fator de aprendizagem quer-se que atualize o próprio Q de maneira moderada.



ARQUITETURA

Criou-se uma classe para gerar o mundo e obter o estado inicial e o estado alvo. Esta classe possui o método de atualizar o estado atual por um novo e o método para mover o agente, fazendo a verificação se o agente colide com algum obstáculo e retorna o próximo estado juntamente com o valor de reforço associado.

Instancia-se o mecanismo de aprendizagem com o conjunto de ações possíveis e começa-se logo por escolher uma ação e obtém-se o estado seguinte com o seu reforço. Depois faz-se o processo de aprendizagem com os dados possuídos, estes são o estado atual, a ação executada, o estado seguinte e o reforço. Internamente faz-se a atualização da memória e a simulação e repete-se todo este processo até o agente atingir o alvo. Assim que o atingir, o agente volta ao estado inicial.

Após uns testes, verifica-se que o agente começa por andar aleatoriamente, mas uma vez que atinge o objetivo, fica logo orientado para a zona onde tem de prosseguir, ficando cada vez mais eficiente, isto é, fazendo o menor de iterações possíveis.

## Realização de uma aplicação de teste que deve incluir a visualização gráfica do comportamento do agente

#### Wavefront

A screenshot of a computer

Description automatically generated with medium confidenceA screenshot of a computer

Description automatically generated with low confidenceA F ilustra o mundo após percorrer todos os estados a partir do alvo, atribuindo um valor cada vez mais pequeno. Depois de gerar o mundo, basta fazer com que o estado inicial percorra pelos estados adjacentes que têm maior valor, pois são esses que estão mais próximos do alvo, pelo caminho mais eficiente, como mostra a F.

#### Dyna-Q

A picture containing qr code

Description automatically generatedA picture containing qr code

Description automatically generatedNa F observa-se pelo número de movimentos efetuados que o agente explorou bastante o mundo, atuando aleatoriamente à procura do alvo. Na F, após já ter atingindo o alvo, vê-se que são precisos muito menos movimentos até lá chegar.

# Tema de escolha livre

# Conclusão