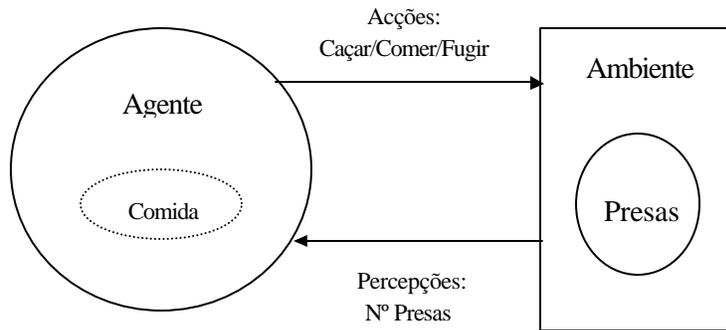


## ALGORITMOS GENÉTICOS/ PROGRAMAÇÃO EVOLUTIVA

Considere que se pretende desenvolver um agente inteligente simples cujo programa corresponda às atitudes de um animal selvagem e interactue com um ambiente (a selva) onde existem presas que podem ser caçadas. O estado interno do agente corresponde à quantidade de comida armazenada (inteiro de 0 a 3). Conforme se ilustra a seguir, a única percepção deste agente é o número de presas que estão ao seu alcance (inteiro de 0 a 3) e em termos de actuações consideramos 3 possibilidades: comer, caçar e fugir (cada uma delas com suas as implicações).



Uma forma de representar a actuação deste agente é codificar os estados possíveis para comida, nº de presas, caçar, comer e fugir em termos de strings binárias:

<b>Comida:</b>	Estados	String
	0	00
	1	01
	2	10
	3	11

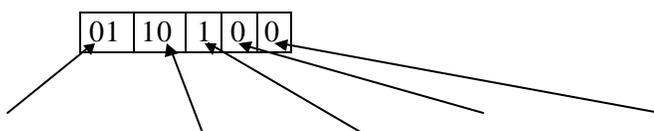
<b>Nº de Presas:</b>	Estados	String
	0	00
	1	01
	2	10
	3	11

<b>Caçar:</b>	Estados	String
	Sim	1
	Não	0

<b>Comer:</b>	Estados	String
	Sim	1
	Não	0

<b>Fugir:</b>	Estados	String
	Sim	1
	Não	0

Nesta perspectiva, cada situação pode ser caracterizada por uma string da forma:



Comida=1      Presas=2      Caçar Não Comer      Não Fugir

Utilizando o método de codificação anteriormente exposto construiu-se o seguinte conjunto de soluções possíveis para o problema do agente:

1100100  
 0011010  
 0101011  
 1111111  
 0000000  
 1011001

Considerando a seguinte função de avaliação

$$f(x) = (\text{comida} + \text{caçar} * (3 - \text{n}^\circ \text{ de presas}) - \text{comer} - \text{fugir}) + 2$$

- Avalie cada uma das soluções.
- Execute uma iteração de um algoritmo genético aplicado a este conjunto de soluções, com uma probabilidade de mutação igual a 0.6 e uma probabilidade cruzamento igual a 1, para tal considere os números aleatórios apresentados na tabela seguinte.
- Indique qual a melhor solução encontrada pelo algoritmo genético. Descodifique esta solução.

N		N		N		N		N		N		N		N		N		N	
1	0,1	11	0,1	21	0,8	31	0,8	41	0,5	51	0,6	61	0,8	71	0,1	81	0,5	91	0,1
2	0,7	12	0,5	22	0,1	32	0,6	42	0,1	52	0,4	62	0,7	72	0,9	82	0,8	92	0
3	0,3	13	0,3	23	0,1	33	0,3	43	0,6	53	0,5	63	0,4	73	0,3	83	0,3	93	0,5
4	0,8	14	0,9	24	0,6	34	0,6	44	0,7	54	0,4	64	0,9	74	0,1	84	0,5	94	0,3
5	0,4	15	0,6	25	0,1	35	0,7	45	0,5	55	0,7	65	0,9	75	0,3	85	0,4	95	0,9
6	0,5	16	0,1	26	0,5	36	0,7	46	0,9	56	0,4	66	0,2	76	0,4	86	0,3	96	0,7
7	0,2	17	0,6	27	0,1	37	0,6	47	0,7	57	0,9	67	0	77	0,2	87	0,6	97	0,3
8	0,1	18	0,5	28	0,6	38	0,2	48	0,1	58	0,4	68	0	78	0,6	88	0,7	98	0,5
9	0,4	19	0,3	29	0,2	39	0,6	49	0,5	59	0,6	69	0,1	79	0,1	89	0,6	99	0,1
10	0	20	0,1	30	0,1	40	0,6	50	0,2	60	0,8	70	0,4	80	0,8	90	0	100	0

Numa obra de construção civil, existem 3 recursos (máquinas por exemplo) A, B e C que podem ser alocados a 5 tarefas possíveis 1, 2, 3, 4 e 5. Conhecidos os custos desta alocação (tabela abaixo), pretende-se, utilizando um programa evolutivo, otimizar, isto é conseguir o menor custo, de alocação. Considere para o efeito que a representação de cada uma das soluções terá a forma  $X_n Y_m Z_k$ , em que n, m e k são tarefas e (1 a 5); X, Y e Z correspondem a recursos (A a C). Os operadores de cruzamento e mutação por sua vez também foram adaptados.

*Cruzamento:* o cruzamento faz-se da forma convencional, como nos algoritmos genéticos, mas o ponto de cruzamento terá de coincidir entre n e Y (*sitecross*=2) ou entre m e Z (*sitecross*=5), correspondendo então à seguinte função (*rnd* é um número aleatório):

$$sitecross = \begin{cases} 2 & \text{se } rnd < 0,5 \\ 4 & \text{se } \text{outro caso} \end{cases}$$

*Mutação:* a mutação processa-se também da forma convencional, mas nos recursos segue a seguinte ordem circular  $B \rightarrow A \rightarrow C \rightarrow B$  e nas tarefas  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 1$ .

Custos de alocação

Tarefa \ Recurso	1	2	3	4	5
A	20	30	10	5	40
B	10	5	10	20	30
C	50	10	5	4	30

Para execução deste programa evolutivo considere ainda:

- A população inicial: A2B3C5, B1C1B3, C2A4B3, B5A5C2;
- A função de avaliação:  $f(x) = \frac{1}{S}$ , em que S corresponde à soma dos custos associados aos recursos e tarefas constantes na solução conforme a tabela de custos de alocação. Ex.  $f(B5A5C2) = \frac{1}{30+40+10} = 0.0125$ .
- A seguinte tabela de números aleatórios:

N		N		N		N		N		N		N		N		N		N	
1	0,1	11	0,1	21	0,8	31	0,8	41	0,5	51	0,6	61	0,8	71	0,1	81	0,5	91	0,1
2	0,7	12	0,5	22	0,1	32	0,6	42	0,1	52	0,4	62	0,7	72	0,9	82	0,8	92	0
3	0,3	13	0,3	23	0,1	33	0,3	43	0,6	53	0,5	63	0,4	73	0,3	83	0,3	93	0,5
4	0,8	14	0,9	24	0,6	34	0,6	44	0,7	54	0,4	64	0,9	74	0,1	84	0,5	94	0,3
5	0,4	15	0,6	25	0,1	35	0,7	45	0,5	55	0,7	65	0,9	75	0,3	85	0,4	95	0,9

6	0,5	16	0,1	26	0,5	36	0,7	46	0,9	56	0,4	66	0,2	76	0,4	86	0,3	96	0,7
7	0,2	17	0,6	27	0,1	37	0,6	47	0,7	57	0,9	67	0	77	0,2	87	0,6	97	0,3
8	0,1	18	0,5	28	0,6	38	0,2	48	0,1	58	0,4	68	0	78	0,6	88	0,7	98	0,5
9	0,4	19	0,3	29	0,2	39	0,6	49	0,5	59	0,6	69	0,1	79	0,1	89	0,6	99	0,1
10	0	20	0,1	30	0,1	40	0,6	50	0,2	60	0,8	70	0,4	80	0,8	90	0	100	0

- a) Avalie cada uma das soluções.
- b) Execute uma iteração deste programa evolutivo, com uma probabilidade de mutação igual a 0.6 e uma probabilidade cruzamento igual a 0.8.
- c) Indique qual a melhor solução encontrada. Justifique.

Numa indústria, existem 3 máquinas A, B e C que podem ser alocadas a 5 operações possíveis 1, 2, 3, 4 e 5, para a fabricação de um produto. Conhecidos os tempos de execução das operações de cada máquina e os custos associados (tabelas abaixo), pretende-se, utilizando um programa evolutivo, otimizar, isto é descobrir a combinação entre máquinas e operações que minimize o tempo e custo de produção. Considere para o efeito que a representação de cada uma das soluções terá a forma  $X_n Y_m Z_k$ , em que n, m e k são operações (1 a 5); e X, Y e Z correspondem a máquinas (A a C). Os operadores de cruzamento e mutação por sua vez também foram adaptados:

**Cruzamento** o cruzamento faz-se da forma convencional, como nos algoritmos genéticos, mas o ponto de cruzamento terá de coincidir entre n e Y (*sitcross*=2) ou entre m e Z (*sitcross*=4), correspondendo então à seguinte função (*rnd* é um número aleatório):

$$sitcross = \begin{cases} 2 & \text{se } rnd < 0.5 \\ 4 & \text{se } \textit{outrocaso} \end{cases}$$

**Mutação** a mutação processa-se também da forma convencional, mas nas máquinas segue a seguinte ordem circular  $A \rightarrow B \rightarrow C \rightarrow A$  e nas operações  $5 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow 5$ .

**Tempos das operações**

Operação \ Máquina	1	2	3	4	5
A	20	30	10	5	40
B	10	5	10	20	30
C	50	10	5	4	30

**Custos das Operações**

Operação \ Máquina	1	2	3	4	5
A	5	2	9	2	3
B	6	1	3	5	3
C	7	9	4	5	3

Tomando como exemplo a solução B5A5C2, o tempo e custo de produção seriam:

$$\text{Tempo} = 30+40+10=80$$

$$\text{Custo} = 3+3+9=15$$

Para execução deste programa evolutivo considere:

- A população inicial: C2A4B3, B5A5C2, A2B3C5, B1C1B3;
- Probabilidade de mutação igual a 0.6;
- Probabilidade cruzamento igual a 0.8;
- A seguinte tabela de números aleatórios:

N		N		N		N		N		N		N		N		N		N	
1	0,1	11	0,1	21	0,8	31	0,8	41	0,5	51	0,6	61	0,8	71	0,1	81	0,5	91	0,1
2	0,7	12	0,5	22	0,1	32	0,6	42	0,1	52	0,4	62	0,7	72	0,9	82	0,8	92	0
3	0,3	13	0,3	23	0,1	33	0,3	43	0,6	53	0,5	63	0,4	73	0,3	83	0,3	93	0,5
4	0,8	14	0,9	24	0,6	34	0,6	44	0,7	54	0,4	64	0,9	74	0,1	84	0,5	94	0,3
5	0,4	15	0,6	25	0,1	35	0,7	45	0,5	55	0,7	65	0,9	75	0,3	85	0,4	95	0,9
6	0,5	16	0,1	26	0,5	36	0,7	46	0,9	56	0,4	66	0,2	76	0,4	86	0,3	96	0,7
7	0,2	17	0,6	27	0,1	37	0,6	47	0,7	57	0,9	67	0	77	0,2	87	0,6	97	0,3
8	0,1	18	0,5	28	0,6	38	0,2	48	0,1	58	0,4	68	0	78	0,6	88	0,7	98	0,5
9	0,4	19	0,3	29	0,2	39	0,6	49	0,5	59	0,6	69	0,1	79	0,1	89	0,6	99	0,1
10	0	20	0,1	30	0,1	40	0,6	50	0,2	60	0,8	70	0,4	80	0,8	90	0	100	0

- a) Desenvolva uma função de avaliação capaz de preterir as soluções mais demoradas e dispendiosas em relação às mais rápidas e mais baratas.
- b) Avalie cada uma das soluções utilizando essa função de avaliação.
- c) Execute uma iteração deste programa evolutivo.
- d) Indique qual a melhor solução encontrada. Justifique.

Um dos grandes problemas de decisão associados à gestão de uma grande superfície comercial está relacionado com a localização dos artigos no espaço de venda. Suponha que o espaço de venda tem a forma de uma matriz quadrada de células identificadas pela linha e coluna (2 linhas e 2 colunas), e que a probabilidade  $P_{ij}$  de venda em cada célula é conhecida e traduzida na seguinte matriz:

	0	1
0	80	70
1	100	30

Considere que apenas são comercializados 4 artigos distintos, e que o lucro que cada artigo representa para a empresa é o seguinte,  $L_i$ :

Artigo	Lucro
00	80
01	5
10	20
11	10

Os algoritmos genéticos/programas evolutivos constituem um dos métodos que pode ser utilizados para pesquisar o espaço de soluções possíveis para este problema. Nesta perspectiva:

- Indique uma forma de representar as soluções para este problema.
- Formule uma função de avaliação que oriente o processo no sentido das soluções que maximizam o lucro desta empresa.
- Proponha uma população inicial com 4 soluções.
- Execute uma iteração do algoritmo genético considerando uma probabilidade de mutação igual a 0.6; uma probabilidade de cruzamento igual a 0.8; e a seguinte tabela de números aleatórios:

N		N		N		N		N		N		N		N		N		N	
1	0,1	11	0,1	21	0,8	31	0,8	41	0,5	51	0,6	61	0,8	71	0,1	81	0,5	91	0,1
2	0,7	12	0,5	22	0,1	32	0,6	42	0,1	52	0,4	62	0,7	72	0,9	82	0,8	92	0
3	0,3	13	0,3	23	0,1	33	0,3	43	0,6	53	0,5	63	0,4	73	0,3	83	0,3	93	0,5
4	0,8	14	0,9	24	0,6	34	0,6	44	0,7	54	0,4	64	0,9	74	0,1	84	0,5	94	0,3
5	0,4	15	0,6	25	0,1	35	0,7	45	0,5	55	0,7	65	0,9	75	0,3	85	0,4	95	0,9
6	0,5	16	0,1	26	0,5	36	0,7	46	0,9	56	0,4	66	0,2	76	0,4	86	0,3	96	0,7
7	0,2	17	0,6	27	0,1	37	0,6	47	0,7	57	0,9	67	0	77	0,2	87	0,6	97	0,3
8	0,1	18	0,5	28	0,6	38	0,2	48	0,1	58	0,4	68	0	78	0,6	88	0,7	98	0,5
9	0,4	19	0,3	29	0,2	39	0,6	49	0,5	59	0,6	69	0,1	79	0,1	89	0,6	99	0,1
10	0	20	0,1	30	0,1	40	0,6	50	0,2	60	0,8	70	0,4	80	0,8	90	0	100	0

## SISTEMAS DE CLASSIFICAÇÃO

Considere que um Sistema de Classificação contém os seguintes classificadores:

Nº	Condição	Acção	Tipo	Valor	Especificidade
1	##0011	010101	ambiente	10	
2	110011	101010	ambiente	5	
3	0#0#1#	111000	ambiente	6	
4	111###	111000	ambiente	2	
5	1100##	001100	ambiente	3	
6	1010#0	110011	ambiente	10	
7	00##11	110011	ambiente	2	
8	111111	110011	ambiente	4	

- Calcule a especificidade de cada classificador
- Indique qual será a acção proposta por este sistema ao ambiente, se a mensagem de entrada for: 110011. Apresente os vários passos do sistema de classificação.
- Calcule o valor dos classificadores após a cobrança das taxas de sobrevivência e de licitação.

Considere os seguintes parâmetros e valores de ruído:

Parâmetro	Valor
n - nº classificadores	8
l - comprimento das mensagens	6
k - tamanho da lista de mensagens	1
$\rho$ - recompensa do ambiente	1
$\beta$ - parte do valor para licitação	0.2
Tvida - taxa de sobrevivência	0.3
Tlanço - taxa de licitação	0.2
Tag - periodicidade do algoritmo genético	1
pm - probabilidade de mutação	0.3
pc - probabilidade de cruzamento	1
npurif - tamanho da população de purificação	3
fpurif - factor de purificação	3
$\sigma$ - desvio padrão para o ruído	0.075

Valores de ruído  $n(\sigma)$  para  $\sigma=0.075$

-0.085      -0.095      0.059      0.039  
 -0.022      0.105      -0.069      -0.200

Em relação aos classificadores do Sistema de Classificação do grupo anterior (I), indique o resultado da execução de um ciclo do Algoritmo Genético.

Considere os seguintes números gerados aleatoriamente:

N		N		N		N		N		N		N		N		N		N	
1	0,1	11	0,1	21	0,8	31	0,8	41	0,5	51	0,6	61	0,8	71	0,1	81	0,5	91	0,1
2	0,7	12	0,5	22	0,1	32	0,6	42	0,1	52	0,4	62	0,7	72	0,9	82	0,8	92	0
3	0,3	13	0,3	23	0,1	33	0,3	43	0,6	53	0,5	63	0,4	73	0,3	83	0,3	93	0,5
4	0,8	14	0,9	24	0,6	34	0,6	44	0,7	54	0,4	64	0,9	74	0,1	84	0,5	94	0,3
5	0,4	15	0,6	25	0,1	35	0,7	45	0,5	55	0,7	65	0,9	75	0,3	85	0,4	95	0,9
6	0,5	16	0,1	26	0,5	36	0,7	46	0,9	56	0,4	66	0,2	76	0,4	86	0,3	96	0,7
7	0,2	17	0,6	27	0,1	37	0,6	47	0,7	57	0,9	67	0	77	0,2	87	0,6	97	0,3
8	0,1	18	0,5	28	0,6	38	0,2	48	0,1	58	0,4	68	0	78	0,6	88	0,7	98	0,5
9	0,4	19	0,3	29	0,2	39	0,6	49	0,5	59	0,6	69	0,1	79	0,1	89	0,6	99	0,1
10	0	20	0,1	30	0,1	40	0,6	50	0,2	60	0,8	70	0,4	80	0,8	90	0	100	0

Apresente a execução dos passos do algoritmo.

## REDES NEURONAIS

1. Representar as gates lógicas or, and e not através de neurónios artificiais utilizando a função de activação step:

a) com t;

b) sem t.

2. Representar através de uma rede neuronal utilizando as unidades anteriores:

a)  $\neg(x \wedge y \vee (\neg x))$

b)  $x \vee y \wedge \neg z \wedge \neg k \vee (x \wedge \neg y)$

atribuir valores às variáveis e testar os resultados.

3. Verificar se a função XOR é linearmene separável. Se não for pode ser implementada através de perceptrons?

4. Treinar um perceptron para os seguintes exemplos. Desenhar a rede.

Entradas			Saída
x1	x2	x3	
1	1	1	1
0	0	0	0
0	1	1	1
0	0	1	1
1	0	1	1

Utilizar o algoritmo de aprendizagem por incremento fixo ( $\eta=0.35$ ) com pesos iniciais todos iguais a 0.5 e executar uma época.

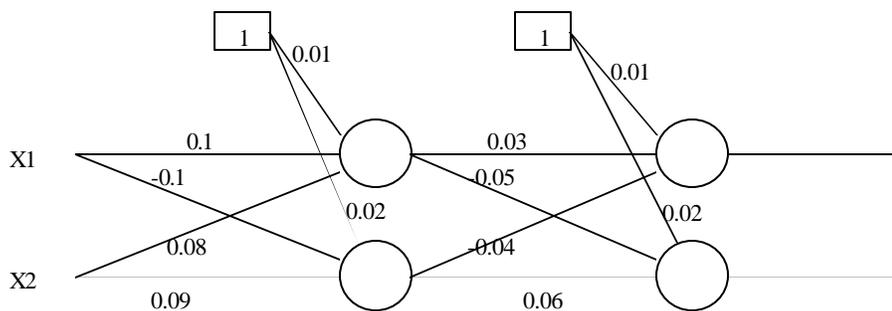
5. Considerar o seguinte conjunto de treino (exemplos) para um perceptron. Cada exemplo tem 6 entradas e uma saída.

	E1	1 1 1 1 1 1 1 0 0 0 0 0 0 0
	E2	0 0 0 1 1 0 0 1 1 0 1 0 1 1
	E3	1 1 1 0 1 0 0 1 1 0 0 0 1 1
Entradas	E4	0 1 0 0 1 0 0 1 0 1 1 1 1 0 1
	E5	0 0 1 1 0 1 1 0 1 1 0 0 1 0
	E6	0 0 0 1 0 1 0 1 1 0 1 1 1 0
Saída	S	1 1 1 1 1 1 1 0 1 0 0 0 0 0 0

Executar o algoritmo de aprendizagem por incremento fixo  $\eta=0.35$ . Determinar o valor dos pesos  $W_0 \dots W_6$ , começando com:

$$W_0=W_1=\dots=W_6=0.0$$

6. Considerar a seguinte rede neuronal .



Executar o algoritmo de aprendizagem por Backpropagation ( $\eta = 0.35$ ) durante uma época. Para tal utilizar o seguinte conjunto de exemplos:

x1	x2	y1	y2
0	0	0.1	0.2
0	1	0.9	0.8
1	1	0.9	0.6
1	0	0.1	0.3

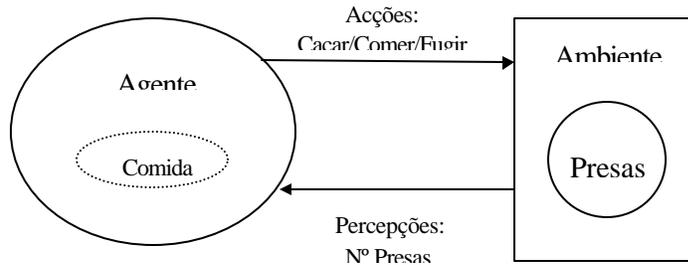
Considere o problema de classificar veículos automóveis nas classes ligeiro e pesado, a partir da tara (peso do veículo) e do número de lugares que este possui. Como não se conhecem as regras para efectuar tal classificação, pretende-se aprender estas regras a partir de um conjunto de exemplos de classificação já efectuados por um perito:

<b>Nº Exemplo</b>	<b>Tara</b>	<b>Nº Lugares</b>	<b>Tipo de Veículo</b>
1	até 1000 kg	2	pesado
2	mais de 1000 kg	7	ligeiro
3	até 1000 kg	5	ligeiro
4	até 1000 kg	2	pesado
5	mais de 1000 kg	5	ligeiro
6	até 1000 kg	5	ligeiro

Uma das possibilidades para adquirir o conjunto de regras para classificação dos veículos é construindo uma rede neuronal que tome como entradas os valores binários dos estados em termos de tara e número de lugares e nas saídas apresente o valor correspondente à classificação efectuada (ligeiro ou pesado). Com esse fim:

- a) Desenhe a rede neuronal que mais se aplica a este problema indicando os nodos, os ramos, as funções de activação e inicialize os pesos associados aos ramos a 0.1.
- b) Indique o valor dos pesos associados à última camada após processar uma época. Considere uma taxa de aprendizagem igual a 0.2.

Considere que se pretende desenvolver um agente inteligente simples cujo programa corresponda às atitudes de um animal selvagem e interactue com um ambiente (a selva) onde existem presas que podem ser caçadas. O estado interno do agente corresponde à quantidade de comida armazenada (inteiro de 0 a 3). Conforme se ilustra a seguir, a única percepção deste agente é o número de presas que estão ao seu alcance (inteiro de 0 a 3) e em termos de actuações consideramos 3 possibilidades: comer, caçar e fugir (cada uma delas com suas as implicações).



Uma forma de representar a actuação deste agente é codificar os estados possíveis para comida, nº de presas, caçar, comer e fugir em termos de strings binárias:

<b>Comida:</b>	Estados	String
	0	00
	1	01
	2	10
	3	11

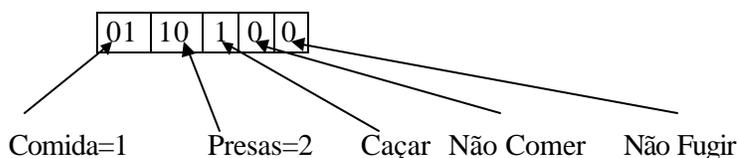
<b>Nº de Presas:</b>	Estados	String
	0	00
	1	01
	2	10
	3	11

<b>Caçar:</b>	Estados	String
	Sim	1
	Não	0

<b>Comer:</b>	Estados	String
	Sim	1
	Não	0

<b>Fugir:</b>	Estados	String
	Sim	1
	Não	0

Nesta perspectiva, cada situação pode ser caracterizada da por uma string da forma:

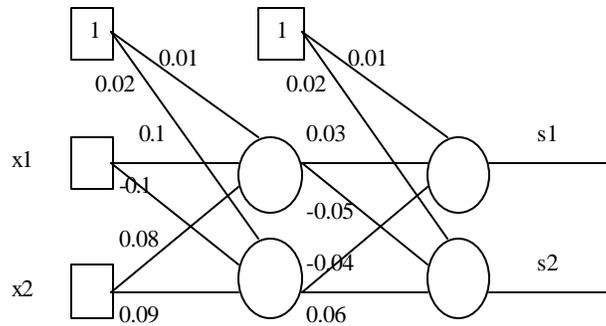


Uma das possibilidades para materializar o programa do agente apresentado é construindo uma rede neuronal que tome como entradas os valores binários dos estados em termos de comida e do nº de presas e nas saídas apresente os valores correspondentes às acções a serem executadas (caçar, comer e fugir). Com esse fim:

- Desenhe a rede neuronal que mais se aplica a este problema indicando os nodos os ramos as funções de activação e inicialize os pesos associados aos ramos.

Indique o valor dos pesos associados à última camada após processar o seguinte exemplo:  
1100101. Considere uma taxa de aprendizagem igual a 0.2.

Considere a seguinte rede neuronal com duas entradas ( $x_1$  e  $x_2$ ), duas saídas ( $s_1$  e  $s_2$ ), estruturada em três camadas de dois *neurónios* cada:

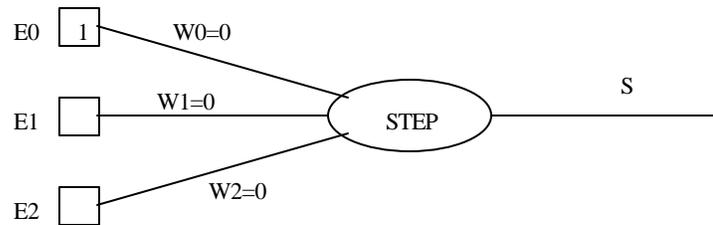


- Se a cada *neurónio* estiver associada a função de activação sigmóide, determine o vector de saída  $\{s_1, s_2\}$  para o vector de entrada  $\{0, 1\}$ .
- Calcule os valores dos erros das unidades na camada de saída, utilizando o algoritmo de backpropagation e considerando o vector de saída calculado na alínea anterior e o vector de saída esperado  $\{1, 1\}$ .
- Calcule os erros propagados à camada anterior.

Desenhe uma rede neuronal que imite o comportamento de um circuito lógico traduzido pela seguinte função:

$$f(x,y,z) = ((x \wedge y) \vee z) \wedge \neg x$$

Considere a seguinte rede neuronal com duas entradas (E1 e E2) e uma saída (S):



- a) Aplique o algoritmo de aprendizagem por incremento fixo ( $\eta=0.3$ ) durante uma época para os seguintes exemplos:

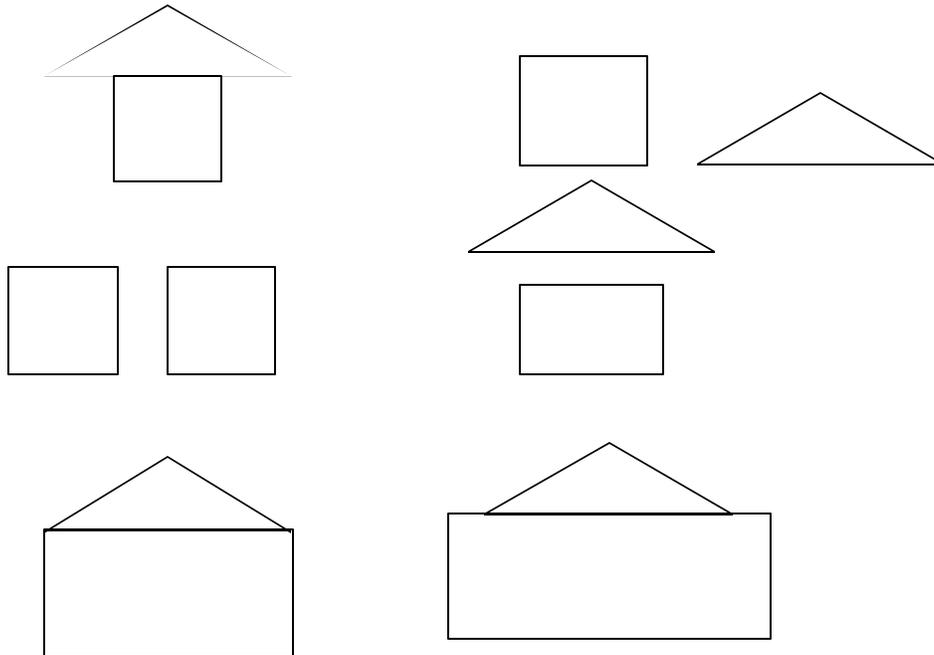
E1	1	0	0	1
E2	1	1	0	0
S	1	0	0	1

determine os valores finais para W0, W1 e W2.

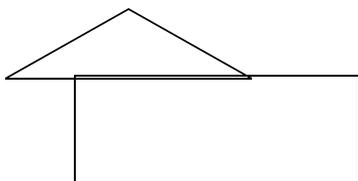
- b) Após o treino efectuado indique qual será a saída apresentada por esta rede para a entrada E1=1, E2=0.

## APRENDIZAGEM DE CONCEITOS ATRAVÉS DE EXEMPLOS

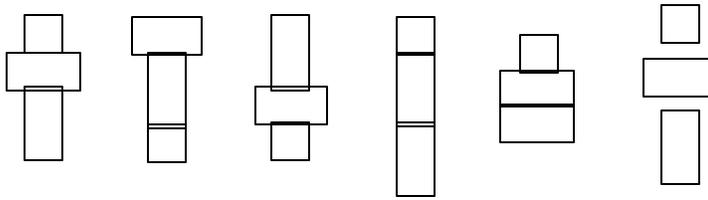
Considerar a sequência de exemplos para o conceito de casa (como uma composição de um triângulo ou um trapézio em cima de um retângulo), a ser submetido a um sistema de aprendizagem de conceitos por indução através de exemplos.



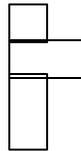
- Representar estes exemplos através de redes semânticas.
- Desenvolver duas regras de indução por forma a que o sistema seja capaz de induzir o conceito de casa.
- Processe os exemplos, usando essas regras, e gere o conceito de casa através da geração de hipóteses.
- Verificar se o caso seguinte unifica o conceito adquirido:



Considere a sequência de exemplos para o conceito de cruz (como uma composição de quadrados e rectângulos), a ser submetido a um sistema de aprendizagem de conceitos por indução através de exemplos.



- Indique qual o universo sobre o qual estamos a actuar e quais os objectos nele contidos.
- Represente estes exemplos através de redes semânticas.
- Indique uma forma de representar os exemplos, as hipóteses e o conceito utilizando lógica de predicados.
- Desenvolva no máximo duas regras de indução por forma a que o sistema seja capaz de induzir o conceito de cruz.
- Processe os exemplos, usando essas regras, e gere o conceito de cruz através da geração de hipóteses. Indique as hipóteses intermédias e o conceito aprendido.
- Verifique se o caso seguinte unifica com o conceito adquirido:



## INDUÇÃO DE ÁRVORES DE DECISÃO

Considere o seguinte conjunto de exemplos, relativos à classificação de alunos segundo os seguintes atributos e respectivos valores

Média:	(A)lta	B(aixa)	
Trabalhos	0	1	2

Aluno	Média	Trabalhos
Aplicado	A	2
Baldas	B	0
Aplicado	A	1
Marrão	A	3
Baldas	B	1
Aplicado	A	0

Através do cálculo dos índices de Gini, induza a árvore de decisão que permita classificar os alunos definidos através dos atributos acima definidos, numa das seguintes classes: Aplicado, Baldas e Marrão.

Apresente os índices de Gini e a árvore induzida.

Considere a seguinte base de dados com informação acerca da triagem dos pacientes na urgência de um hospital.

Sintoma	Idade	Sexo	Unidade de tratamento
Febre	1 a 40	M	A
Febre	mais de 40	F	B
Febre	mais de 40	M	A
Febre	1 a 40	F	B
Febre	1 a 40	F	A
Febre	1 a 40	M	A
Hemorragia	1 a 40	F	B
Hemorragia	mais de 40	M	C
Hemorragia	1 a 40	M	C
Queimaduras	mais de 40	M	C
Queimaduras	1 a 40	M	B
Queimaduras	1 a 40	F	B
Queimaduras	mais de 40	M	C
Queimaduras	mais de 40	F	A

- a) Induza a árvore de decisão que permite classificar os pacientes segundo as unidades de tratamento disponíveis. Indique todos os cálculos.
- b) Efectue a poda da árvore.

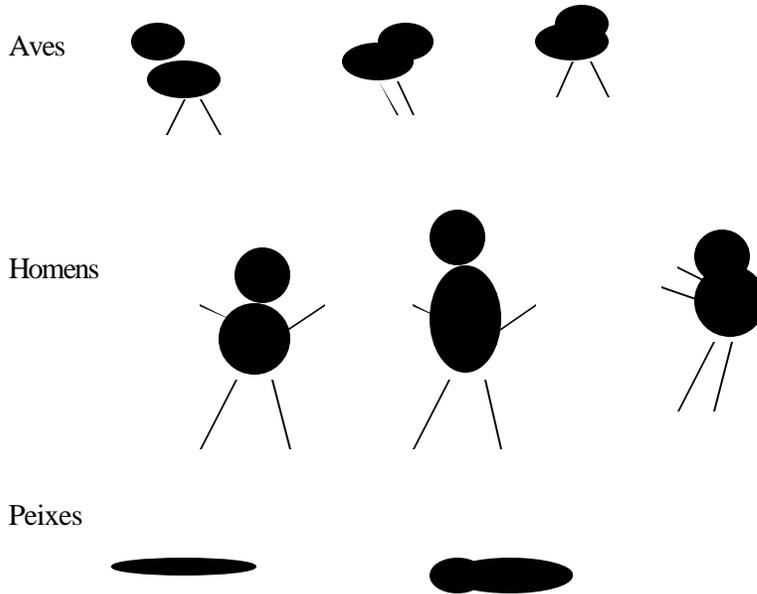
Considere o problema de classificar veículos automóveis nas classes ligeiro e pesado, a partir da tara (peso do veículo) e do número de lugares que este possui. Como não se conhecem as regras para efectuar tal classificação, pretende-se aprender estas regras a partir de um conjunto de exemplos de classificação já efectuados por um perito:

<b>Nº Exemplo</b>	<b>Tara</b>	<b>Nº Lugares</b>	<b>Tipo de Veículo</b>
1	até 1000 kg	2	pesado
2	mais de 1000 kg	7	ligeiro
3	até 1000 kg	5	ligeiro
4	até 1000 kg	2	pesado
5	mais de 1000 kg	5	ligeiro
6	até 1000 kg	5	ligeiro

- a) Induza a árvore de decisão que permite classificar os veículos automóveis segundo os tipos definidos. Indique todos os cálculos.

## APRENDIZAGEM

Uma câmara fotográfica captou as seguintes formas planas (silhuetas) a partir de um conjunto de objectos tridimensionais.



Com base nestes exemplos, pretende-se aprender por indução os conceitos de: ave, peixe e homem. Para caracterizar estes objectos considere os seguintes atributos:

Atributo	Valores possíveis
tamanho	pequeno, grande
extremidades	0, 2, 3, 4
formas ovais	1, 2

### I

- Faça uma representação dos exemplos em termos do conjunto de atributos definido;
- Induza e represente os conceitos utilizando o algoritmo de aprendizagem de indução de conceitos dado nas aulas;
- Indique se o seguinte caso unifica com um dos conceitos e em caso afirmativo identifique o conceito: [tamanho=pequeno, extremidades=0].
- Faça uma representação dos exemplos em termos do conjunto de atributos definido;
- Induza a árvore de decisão que permite classificar os exemplos segundo os conceitos definidos. Indique todos os cálculos.
- Entende que é necessário efectuar a poda da árvore? Justifique.
- Um dos passos executados antes de se aplicar uma ferramenta de *Data-Mining* a uma base de dados é a investigação das relações ou associações entre os atributos. Indique

para este problema quais as associações existentes e o seu peso (ocorrências) entre os atributos e os conceitos. Tome como base os exemplos representados na alínea a).

- e) Se pretendesse construir um sistema de classificação adequado a este problema qual deveria ser a forma dos classificadores em termos de codificação das *condições* e *acções*? Dê um exemplo.

## II

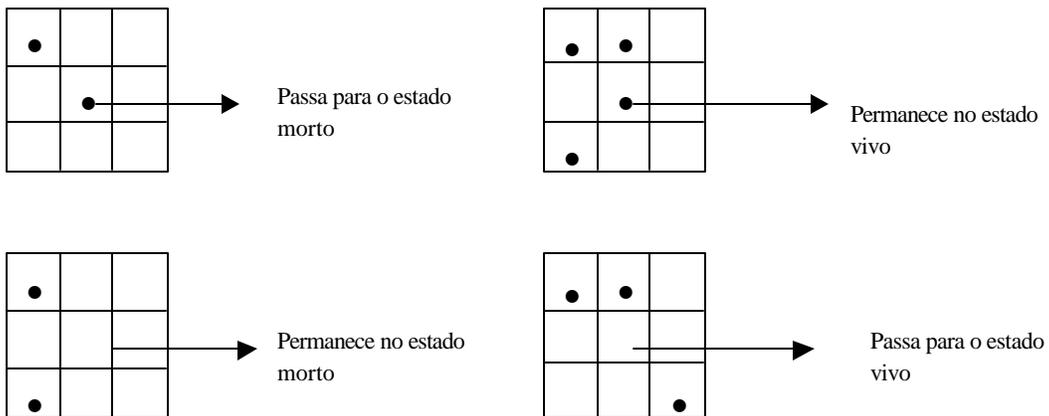
Uma outra possibilidade para adquirir o conjunto de regras para classificação dos exemplos apresentados no grupo I é construindo uma rede neuronal que tome como entradas os valores dos atributos e nas saídas apresente o valor correspondente à classificação efectuada (conceito). Com esse fim:

- a) Desenhe a rede neuronal que mais se aplica a este problema indicando os nodos, os ramos, as funções de activação e inicialize os pesos associados aos ramos com o valor 0.1.
- b) Calcule a classificação efectuada pela rede desenvolvida para o primeiro exemplo (ave).
- c) Indique o valor dos pesos associados à última camada após processar um ciclo de aprendizagem. Considere uma taxa de aprendizagem igual a 0.2.

*VIDA* é um jogo individual que é jogado numa grelha de células quadradas. Num dado instante, uma célula pode estar no estado vivo ou morto. O padrão de comportamento das células é determinado pelas seguintes regras:

1. Se uma célula está viva, continuará viva se estiver rodeada por exactamente 2 ou 3 células vivas. Noutro caso passará ao estado morto.
2. Se uma célula está morta, passará ao estado vivo se estiver rodeada por exactamente 3 células vivas. Noutro caso permanecerá morta.

Considere os quatro exemplos seguintes com uma grelha de 3x3 células e como referência a célula na posição central (as células com • estão vivas, as restantes estão mortas):



Imagine que as regras para o jogo da *VIDA* não eram conhecidas. Mas existia um conjunto de exemplos disponível. Seria possível “aprender” as regras associadas ao jogo através do processamento desses exemplos?

- a) Seleccione o paradigma de aprendizagem que considere ser mais adequado.
- b) Indique uma forma de representar os exemplos. Represente os quatro exemplos apresentados.
- c) Desenvolva o sistema que indicou em a).
- d) Execute um ciclo de treino/aprendizagem no sistema desenvolvido, utilizando os quatro exemplos apresentados.

## DESCOBERTA DE CONHECIMENTO EM BD/ DATA MINING

Considere a seguinte base de dados com informação acerca dos pacientes de um hospital, com a mesma doença, e da sua resposta a duas drogas diferentes experimentadas. Utilize um dos métodos de *Data Mining* estudados para descobrir um padrão capaz de prever qual a droga apropriada para um novo paciente.

- Identifique de entre os objectivos definidos em *Data Mining* aquele que está associado a este problema.
- Defina qual o método mais apropriado para descobrir o padrão pretendido.
- Utilize o método descrito na alínea anterior para gerar o padrão pretendido. Apresente o padrão descoberto.
- Traduza o padrão descoberto para um conjunto de regras *se...então*.
- Indique uma forma de calcular os erros cometidos por este padrão.
- Investigue as relações ou associações entre a *Pressão Sanguínea* e a *Droga*.

Caso	<i>Pressão Sanguínea</i>	<i>Nível de Sódio no Sangue</i>	<i>Droga</i>
1	Alta	Alto	Y
2	Alta	Alto	Y
3	Alta	Baixo	X
4	Baixa	Alto	X
5	Baixa	Baixo	X
6	Baixa	Baixo	X