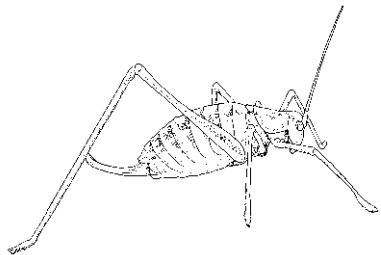


O problema de classificação

(definição informal)

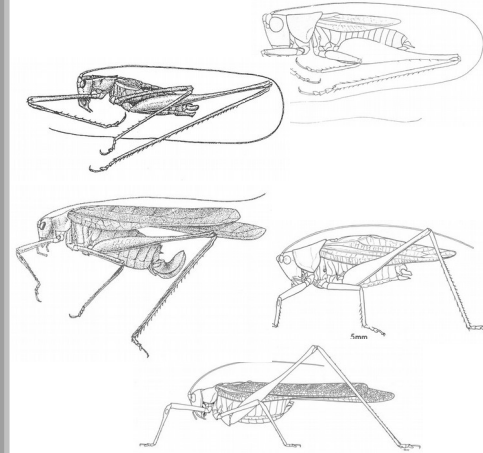
Dada uma coleção de dados detalhados, neste caso 5 exemplos de **Esperança** e 5 do **Gafanhoto**, decida a qual tipo de inseto o exemplo não rotulado pertence.

Obs: **Esperança** : tipo de gafanhoto verde.

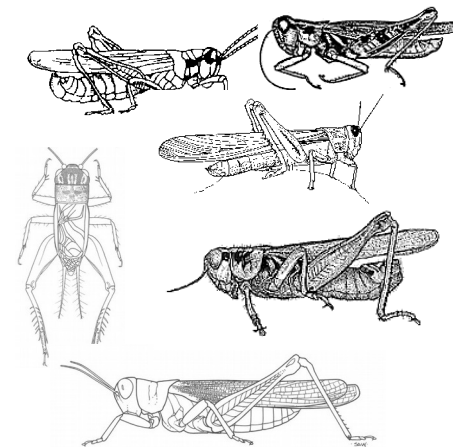


Esperança ou **Gafanhoto**?

Esperança



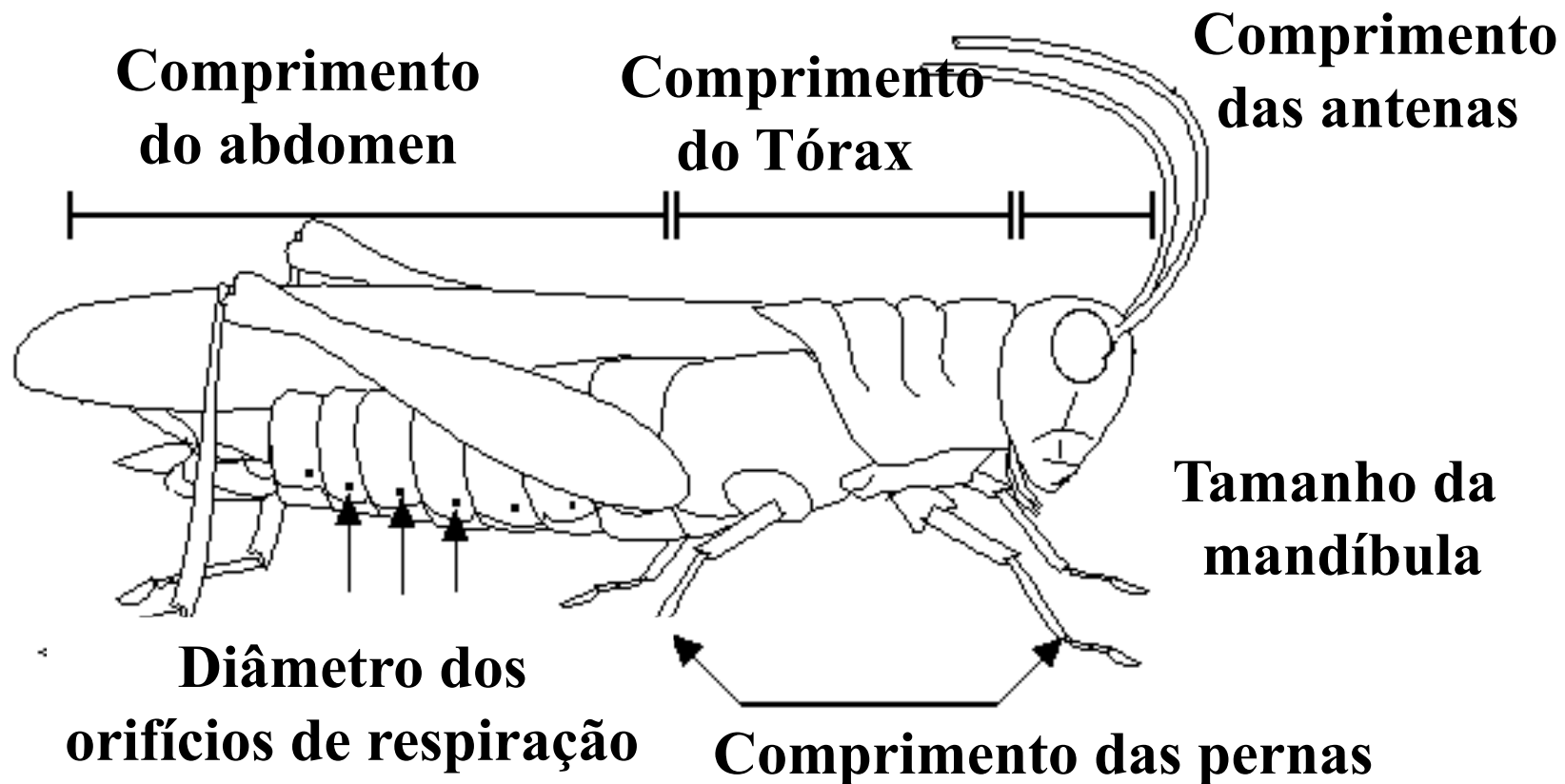
Gafanhoto



Para qualquer domínio de interesse
podemos medir *características*

Cor {Verde, Marrom, Cinza, Outra}

Tem asas?



Podemos armazenar as *características* em bases de dados

O problema de classificação agora pode ser expresso da seguinte forma:

- Dada uma base de treinamento(**Minha_Coleção**), prever o rótulo da **classe dos exemplos ainda não vistos**

Minha_Coleção

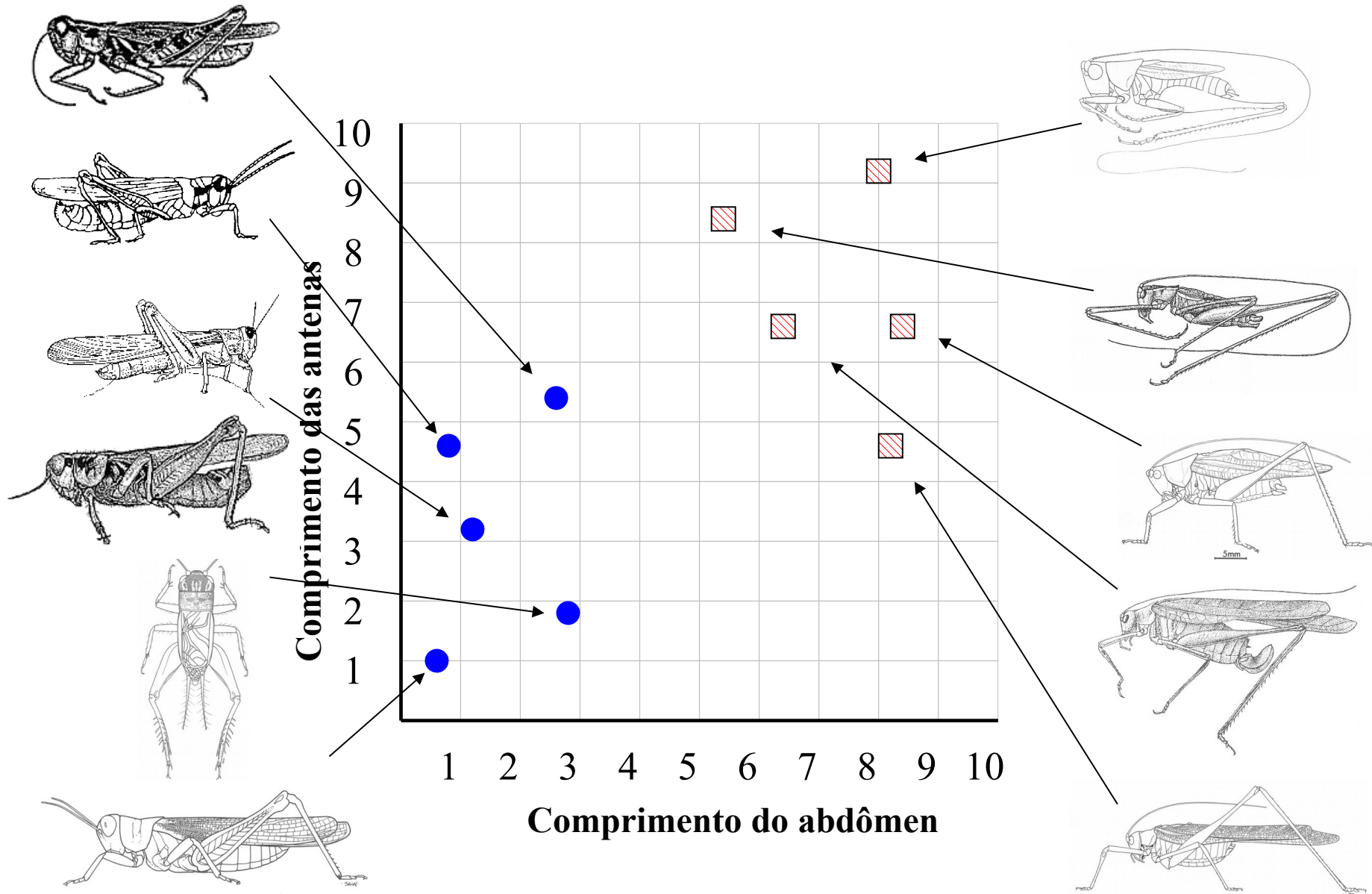
ID do inseto	Comp. do abdômen	Comp. das antenas	Classe do inseto
1	2.7	5.5	Gafanhoto
2	8.0	9.1	Esperança
3	0.9	4.7	Gafanhoto
4	1.1	3.1	Gafanhoto
5	5.4	8.5	Esperança
6	2.9	1.9	Gafanhoto
7	6.1	6.6	Esperança
8	0.5	1.0	Gafanhoto
9	8.3	6.6	Esperança
10	8.1	4.7	Esperança

Exemplo não visto =

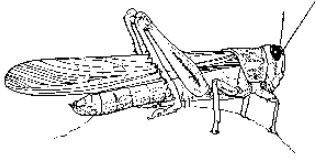
11	5.1	7.0	??????
----	-----	-----	--------

Gafanhoto

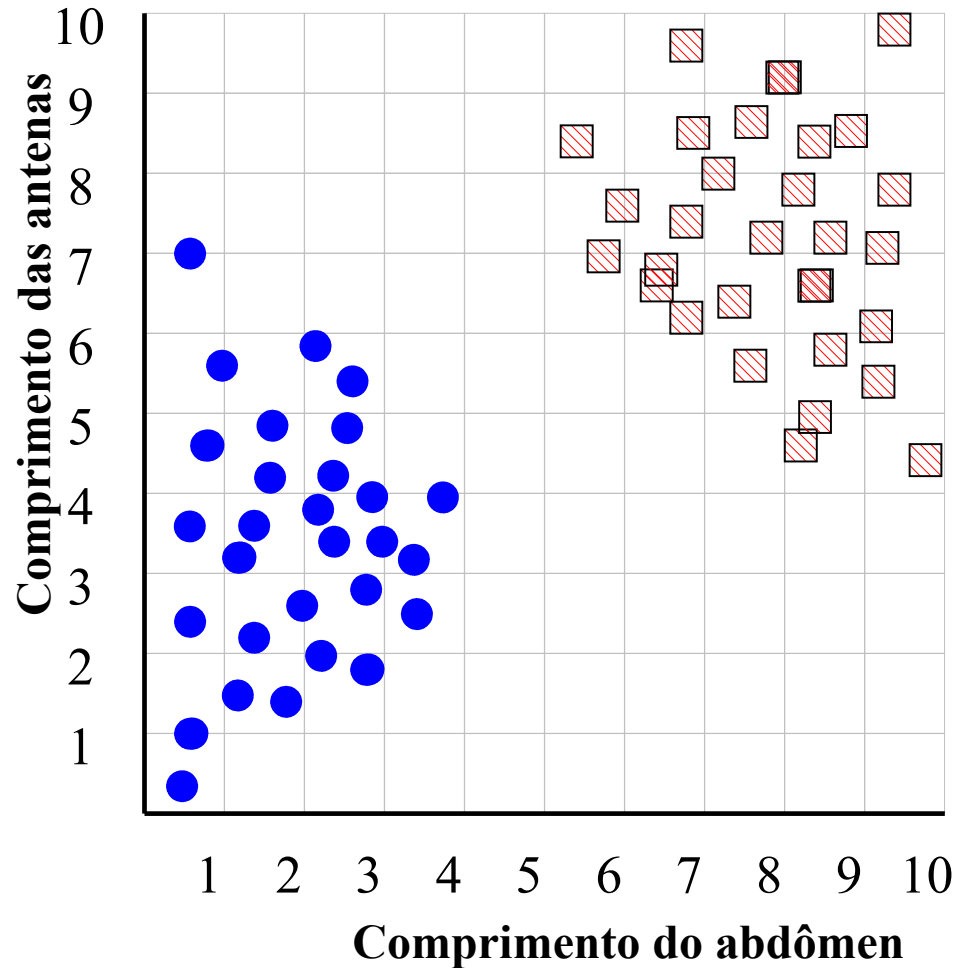
Esperança



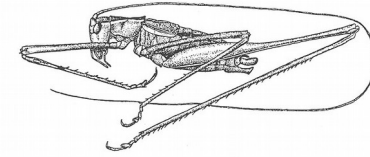
Gafanhoto



Também utilizaremos esta base de dados maior para motivação ...



Esperança

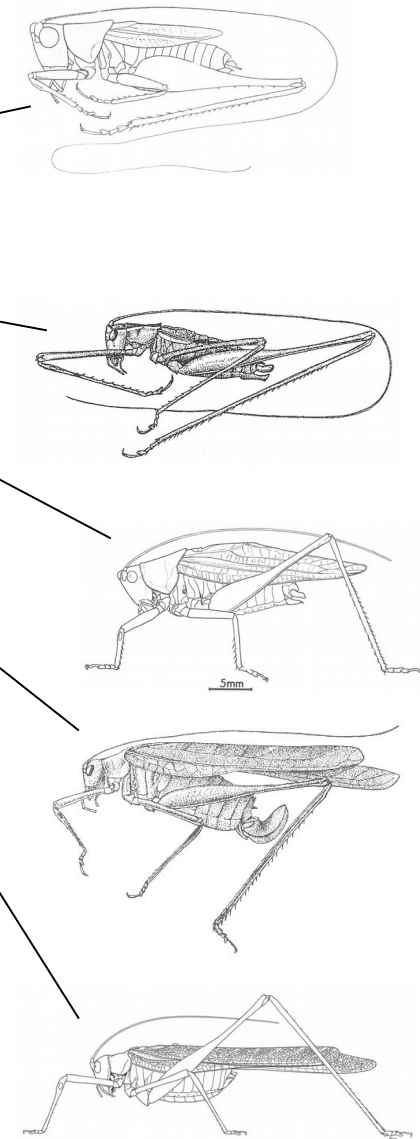
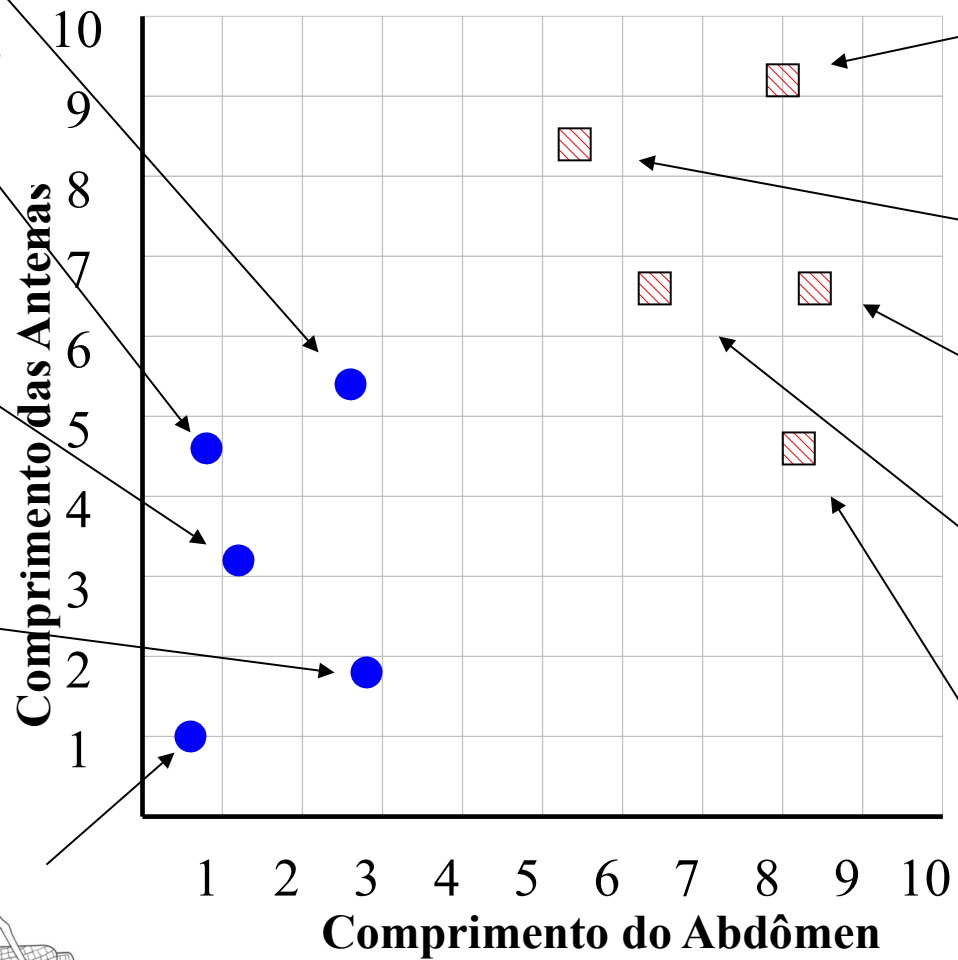
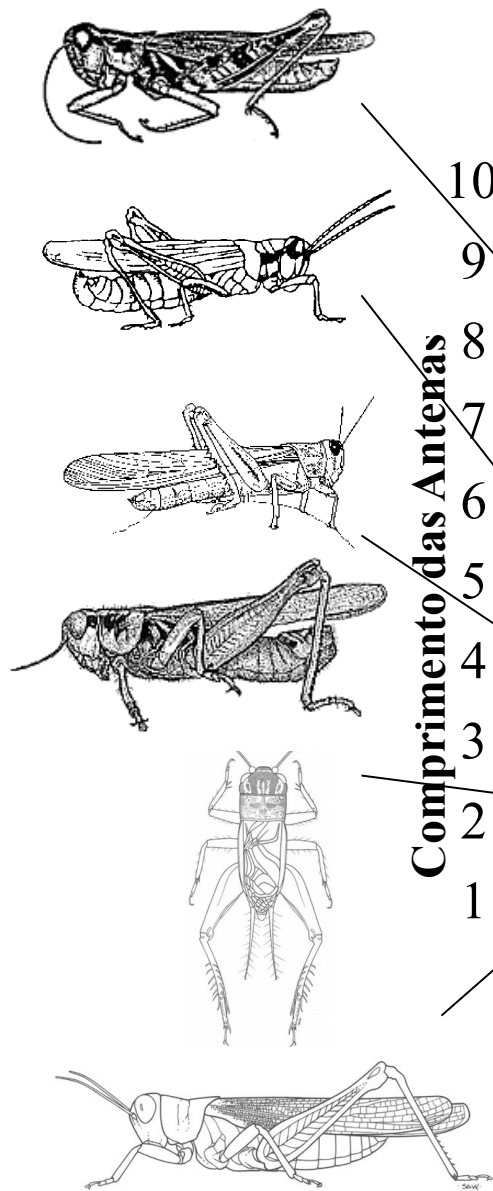


Cada um destes objetos de dados é chamado de...

- exemplar
- exemplo (de treinamento)
- instância
- tupla

Gafanhoto

Esperança



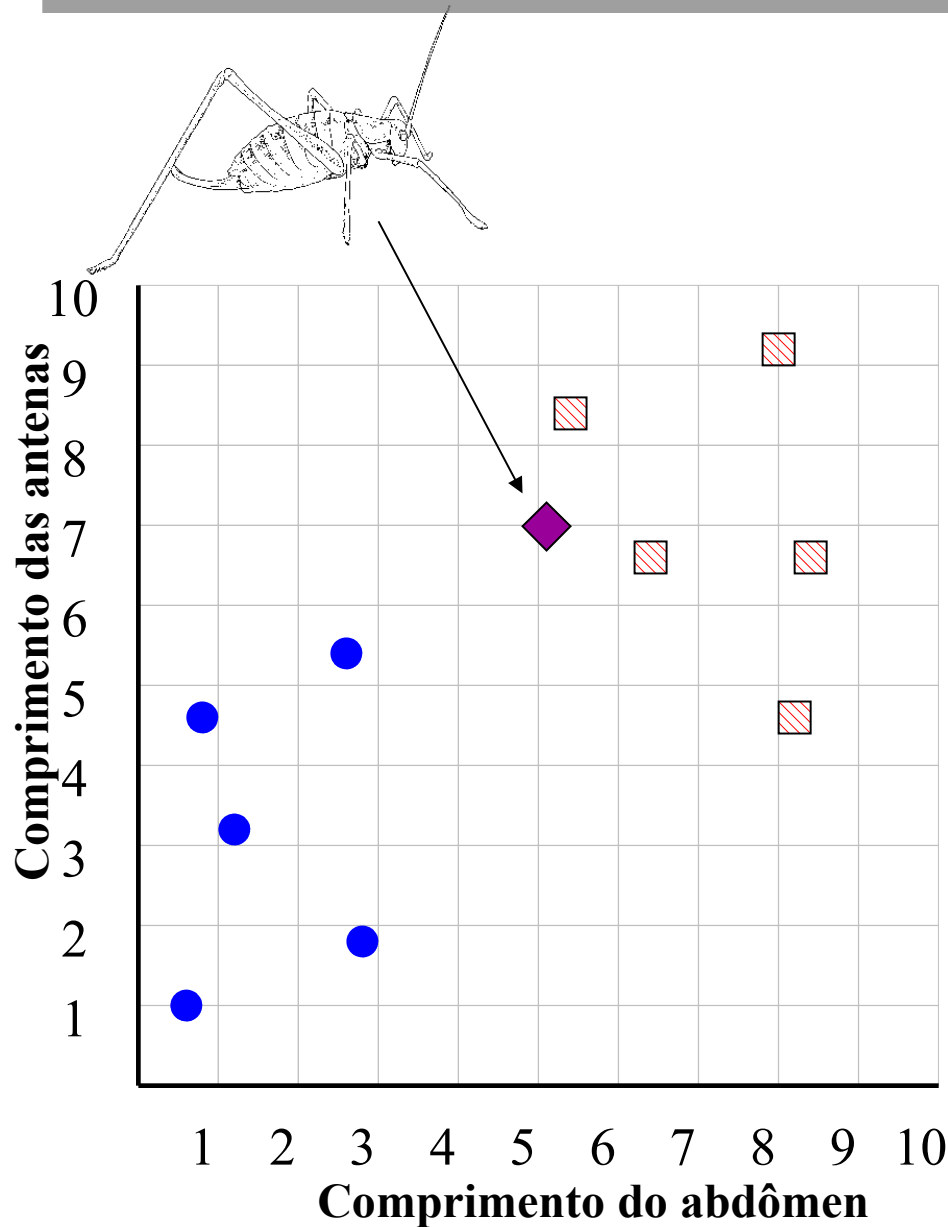
Exemplo não visto antes =

11

5.1

7.0

???????



- Podemos “projetar” o exemplo não visto antes dentro do mesmo espaço que a base de dados.
- Acabamos de abstrair os detalhes do nosso problema particular. Será muito mais fácil conversar sobre pontos no espaço.

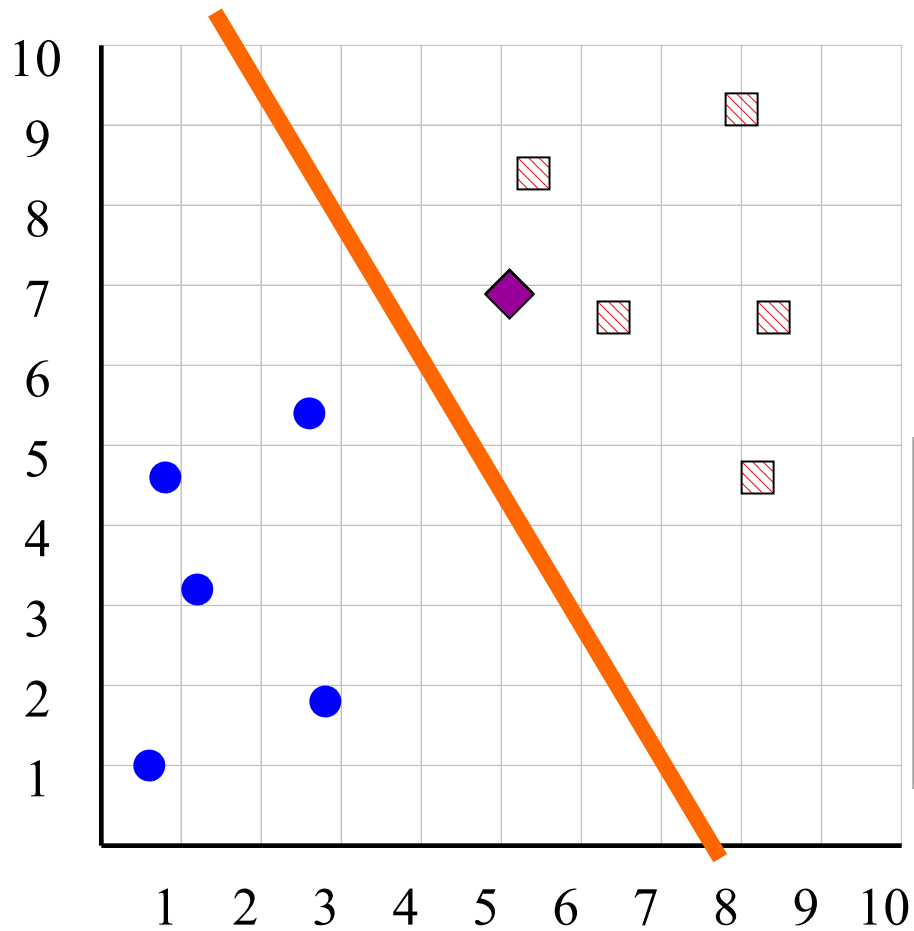
■ **Esperança**

● **Gafanhoto**

Classificador Linear Simples



R.A. Fisher
1890-1962



Se **exemplo não visto antes** está acima da linha

Então

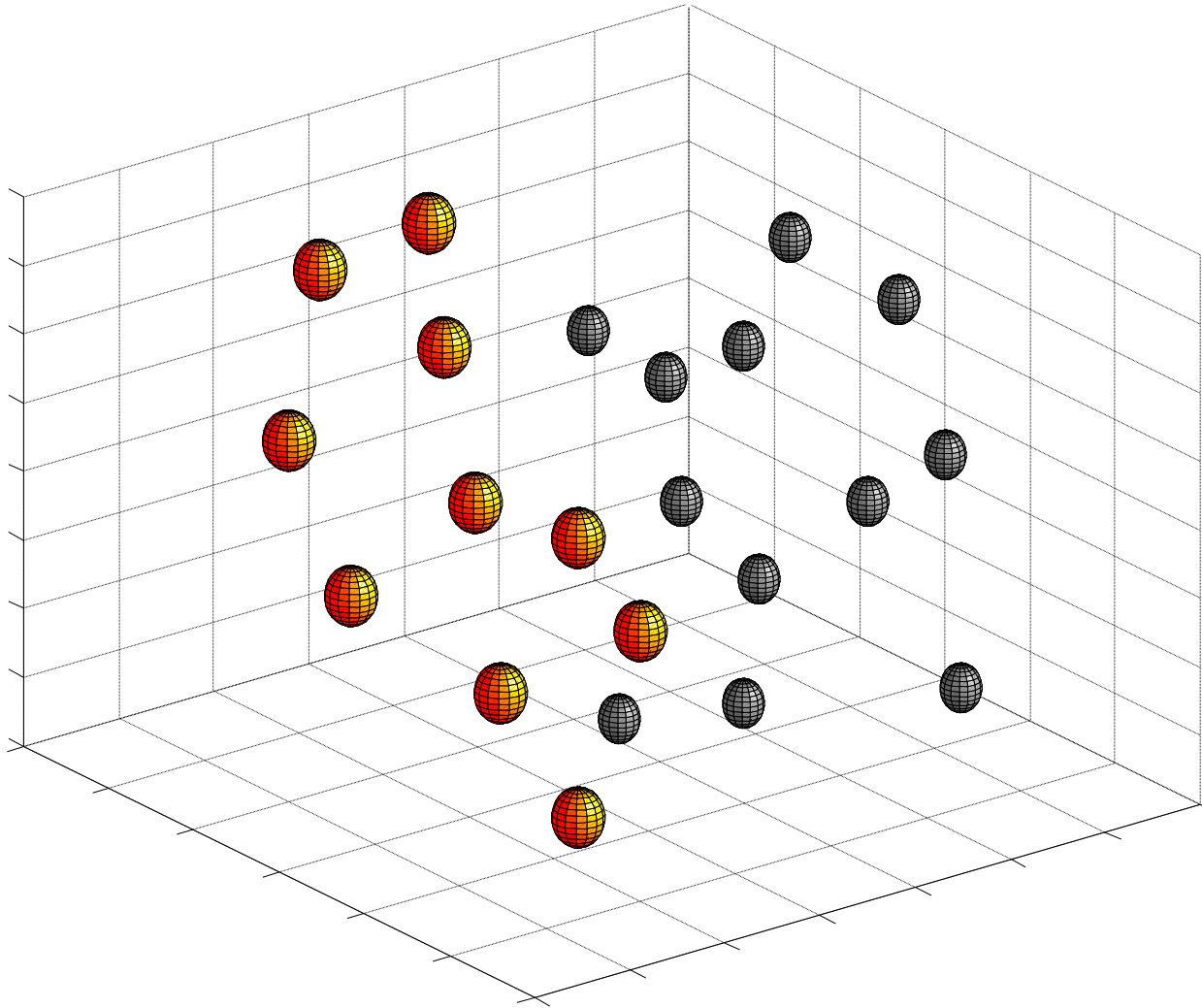
classe é **Esperança**

senão

classe é **Gafanhoto**

■ **Esperança**
● **Gafanhoto**

O classificador linear simples
é definido para espaços dimensionais maiores...



... podemos visualizá-lo
como sendo um
hiperplano
n-dimensional

