



Concepção e desenvolvimento de pequenos robôs para auto-aprendizagem

Hugo Tavares

Sob a orientação científica do Prof. Doutor
Vítor Manuel Ferreira Dos Santos



- Objectivos e Motivação
- Projecto dos robôs
- Software desenvolvido
- Testes e Resultados
- Conclusões e trabalhos futuros

Objectivos e Motivação



- Objectivos e Motivação
 - Projecto dos robôs
 - Software
 - Testes e Resultados
 - Conclusões e Trabalhos futuros
- Projectar, desenvolver e construir pequenos robôs orientados para aprendizagem de comportamentos de locomoção;
 - Implementação de um algoritmo de aprendizagem nos robôs para provar a capacidade de aprendizagem dos mesmos.

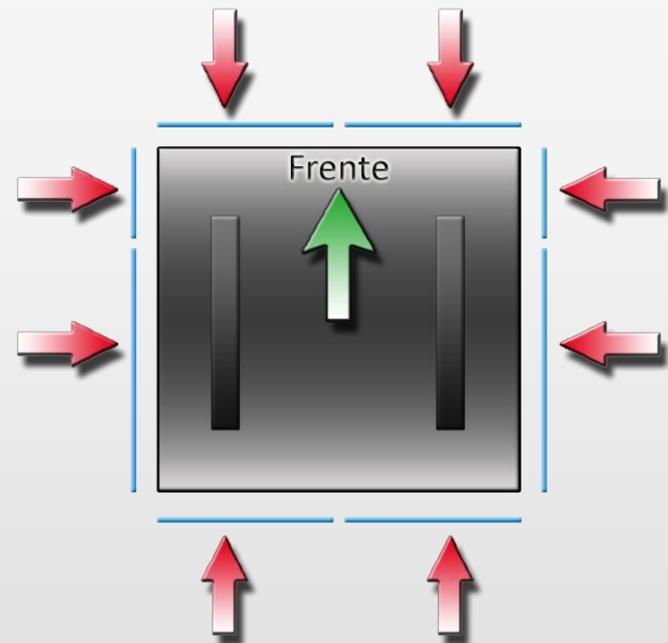


- Objectivos e Motivação
 - Projecto dos robôs
 - Software
 - Testes e Resultados
 - Conclusões e Trabalhos futuros
- SALbot:
 - Chassi;
 - Sistema de pára-choques;
 - Placa de circuito principal;
 - Placa de circuito auxiliar;
 - Sensores de distância:
 - Infra-vermelhos;
 - Ultra-sons.
 - Servos.



- Sistema de Pára-choques:

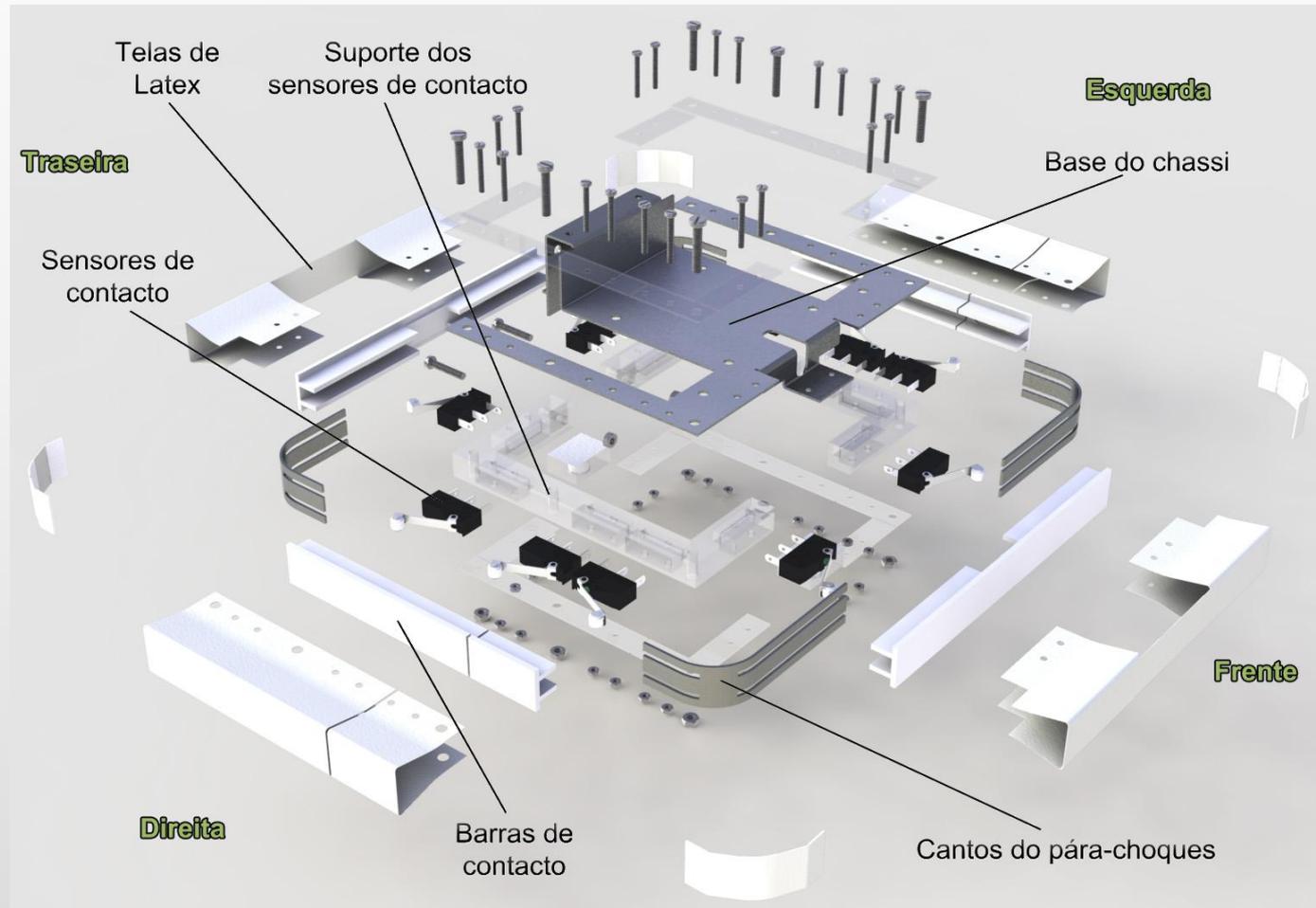
- Fiabilidade;
- Robustez;
- Zonas de contacto;



Projecto dos robôs



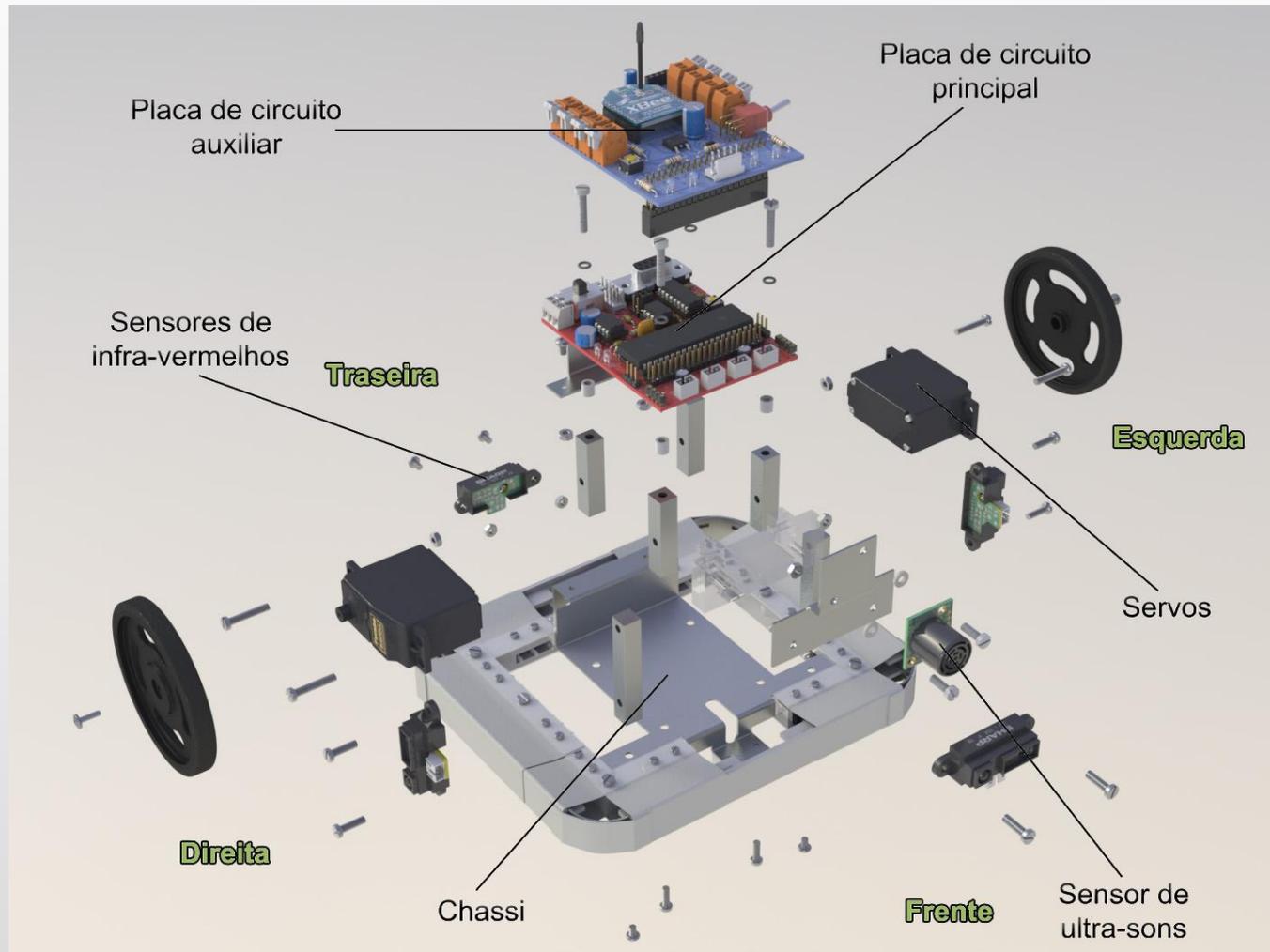
- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros



Projecto dos robôs



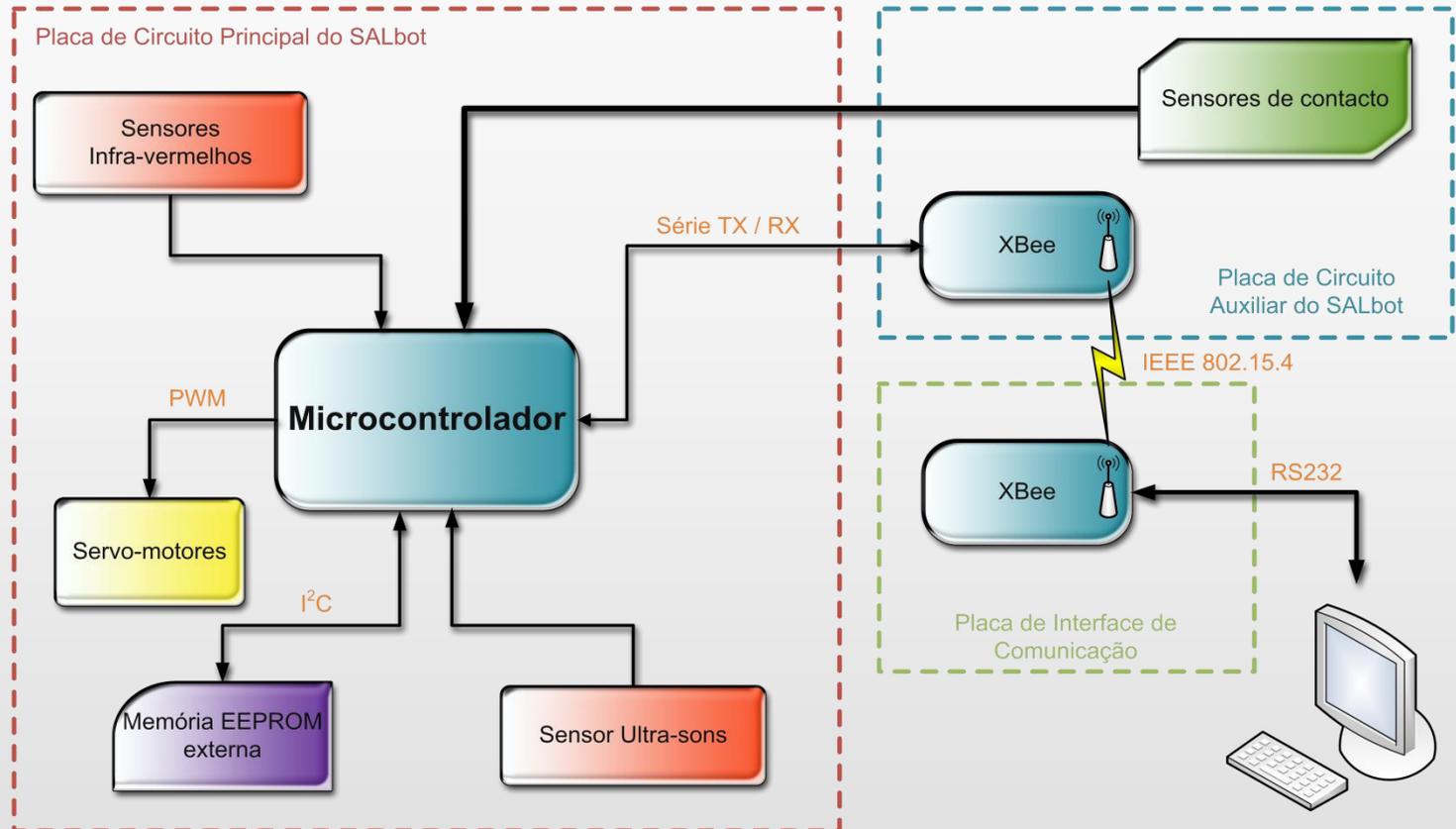
- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros



Projecto dos robôs



- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros





- *Reinforcement learning:*

- Q-learning:

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right]$$

- SARSA:

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha [r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

$a \rightarrow$ acção

$s \rightarrow$ estado

$r \rightarrow$ recompensa

$\alpha \rightarrow$ coeficiente de aprendizagem;

$\gamma \rightarrow$ factor de desconto



- Pseudo-código do algoritmo SARSA:

Inicializar:

$Q(s, a) \leftarrow$ arbitrário

Repetir para cada episódio:

Inicializar s

$a' \leftarrow$ tendo s' , baseado em Q , usando uma política (por ex.: ϵ -greedy)

Repetir para cada passo do episódio até s terminar:

Executar a acção a

Observar o próximo estado s' e a recompensa r

$a' \leftarrow$ tendo s' , baseado em Q , usando uma política (por ex.: ϵ -greedy)

$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma Q(s', a') - Q(s, a)]$

$s \leftarrow s'$

$a \leftarrow a'$

- Objectivos e Motivação
- Projecto dos robôs
- **Software**
- Testes e Resultados
- Conclusões e Trabalhos futuros

Software desenvolvido



- Objectivos e Motivação
- Projecto dos robôs
- **Software**
- Testes e Resultados
- Conclusões e Trabalhos futuros

Tabela de Aprendizagem

$\alpha = 0.5$
 $\gamma = 0.5$

Estados \ Acções	0	1
0	0	0
1	0	0



$$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma Q(s', a') - Q(s, a)]$$

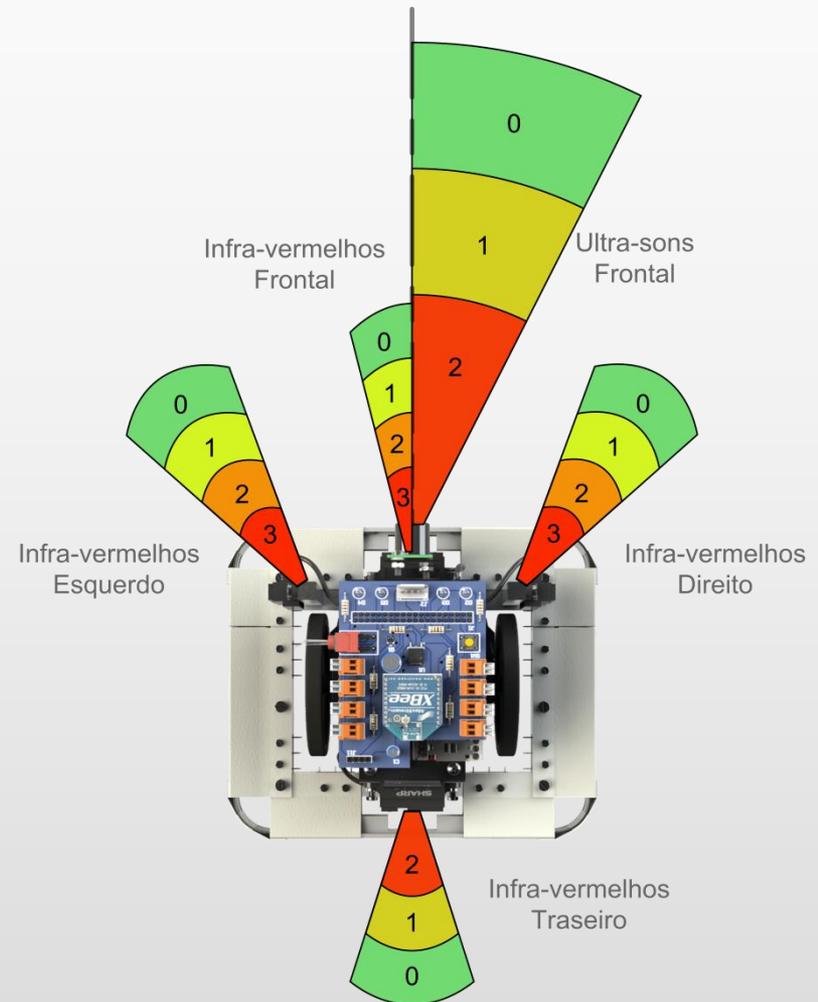
$s = 0$
 $a = 0$

Software desenvolvido



- Estados:

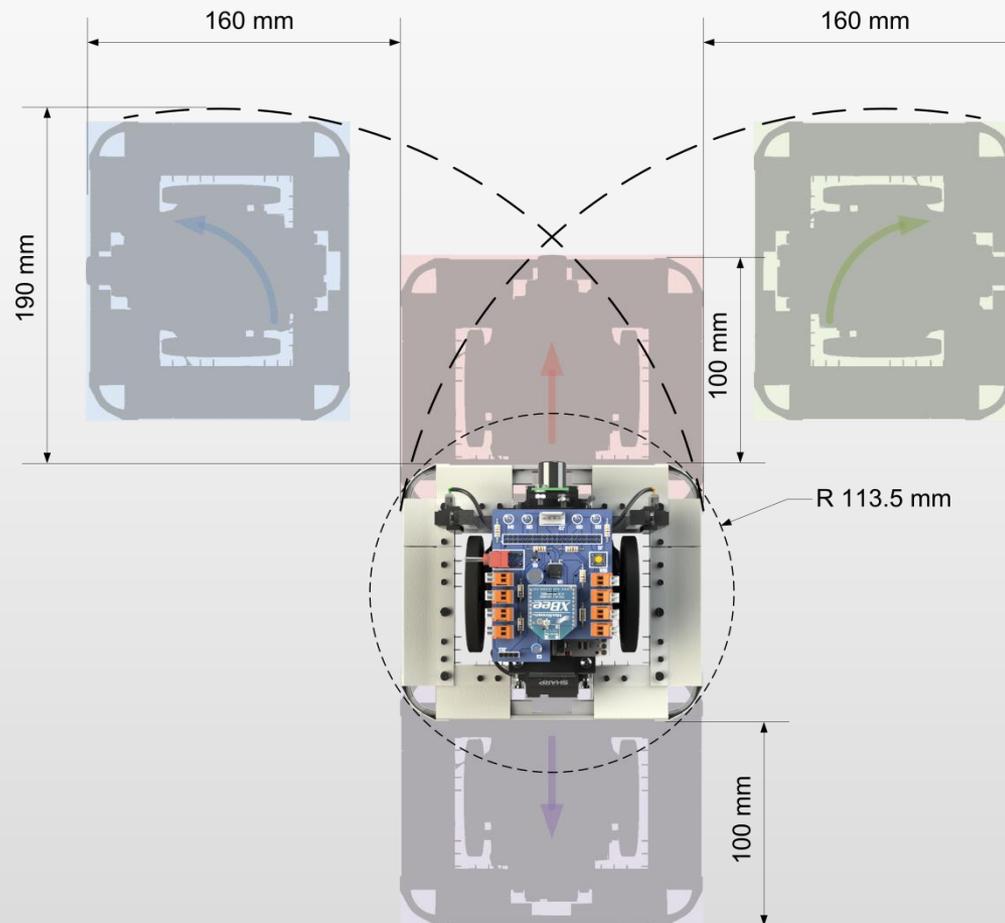
$$4 \times 4 \times 4 \times 3 \times 3 \times 16 = 9216$$



- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros



- Acções (6):



- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

Testes e resultados



- Testes realizados em 3 tipos de ambientes:
 - Pequeno ($1/4$ da área do ambiente de testes);
 - Médio ($1/2$ da área do ambiente de testes);
 - Totalidade do ambiente de testes.

- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

Testes e resultados

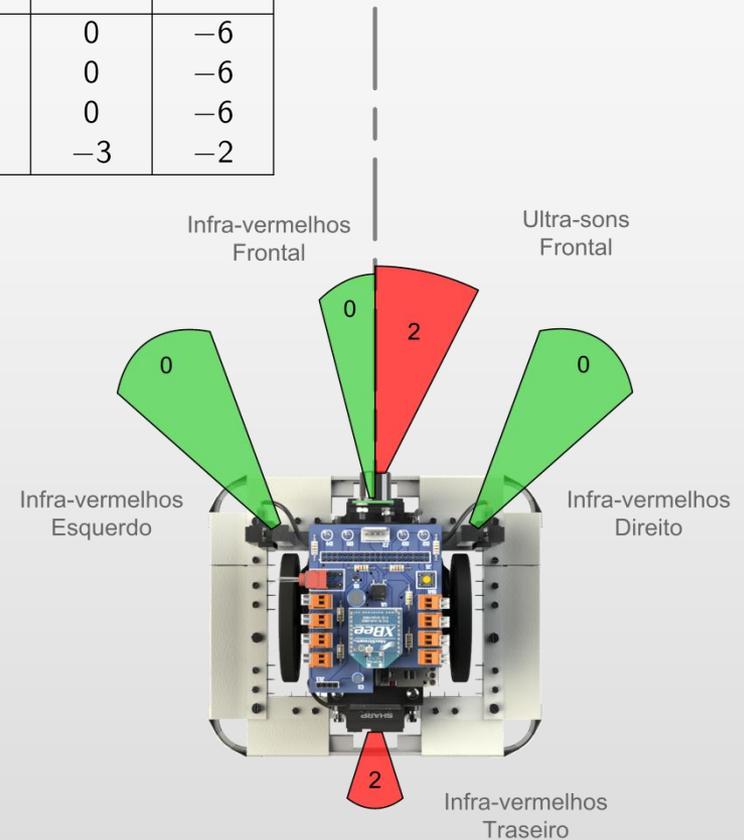


Testes e resultados



- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

		Estado 416					
Iteração \ Acção		0	1	2	3	4	5
500		7	8	0	0	0	-6
1000		7	10	0	0	0	-6
2000		13	7	0	5	0	-6
5000		13	10	8	5	-3	-2

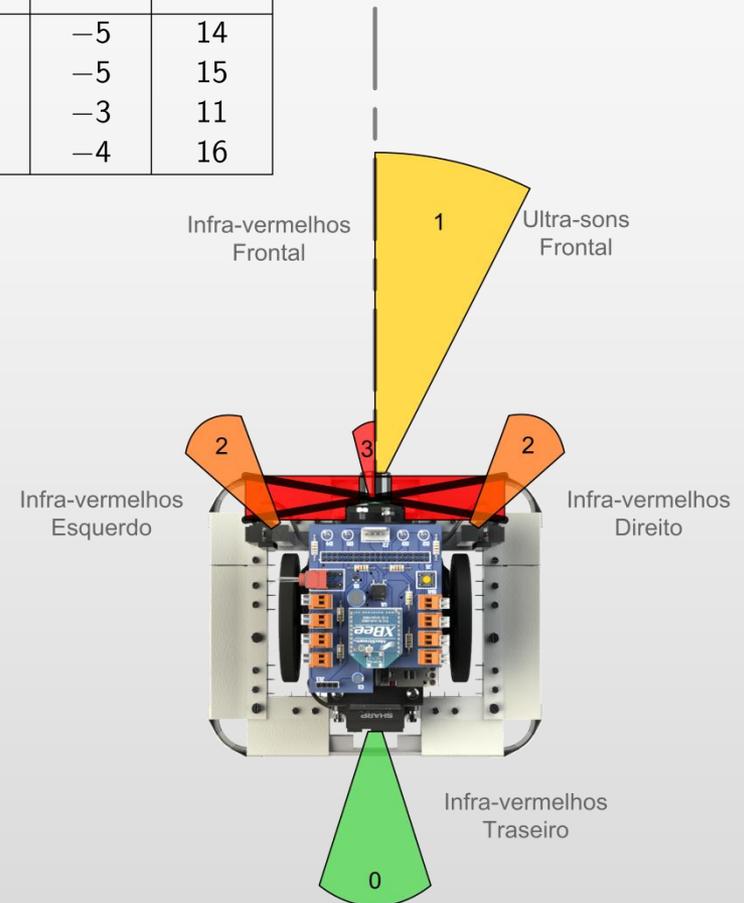


Testes e resultados



- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

		Estado 5921					
Iteração	Acção	0	1	2	3	4	5
500		-11	-4	-3	-9	-5	14
1000		-1	-2	-5	2	-5	15
2000		-3	-4	-7	3	-3	11
5000		-5	-10	-2	-2	-4	16



Testes e resultados



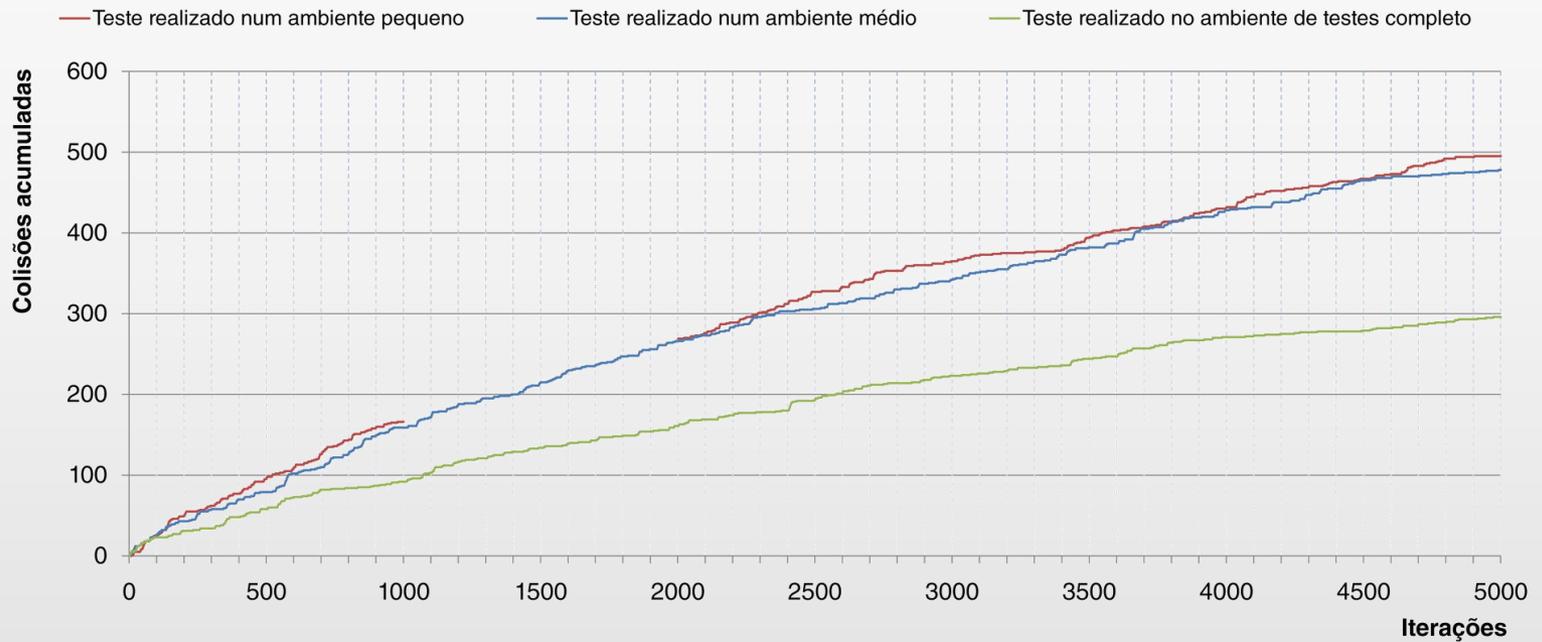
- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

		Estados visitados durante a etapa			Novos estados desde a etapa anterior			Estados inalterados desde a etapa anterior		
		¼	½	1	¼	½	1	¼	½	1
Iteração	Teste	¼	½	1	¼	½	1	¼	½	1
500		140	112	96	–	–	–	–	–	–
1000		192	165	126	52	53	30	46	34	30
2000		245	231	178	53	66	52	65	55	36
5000		345	297	255	100	66	77	67	51	40

Resultados da aprendizagem



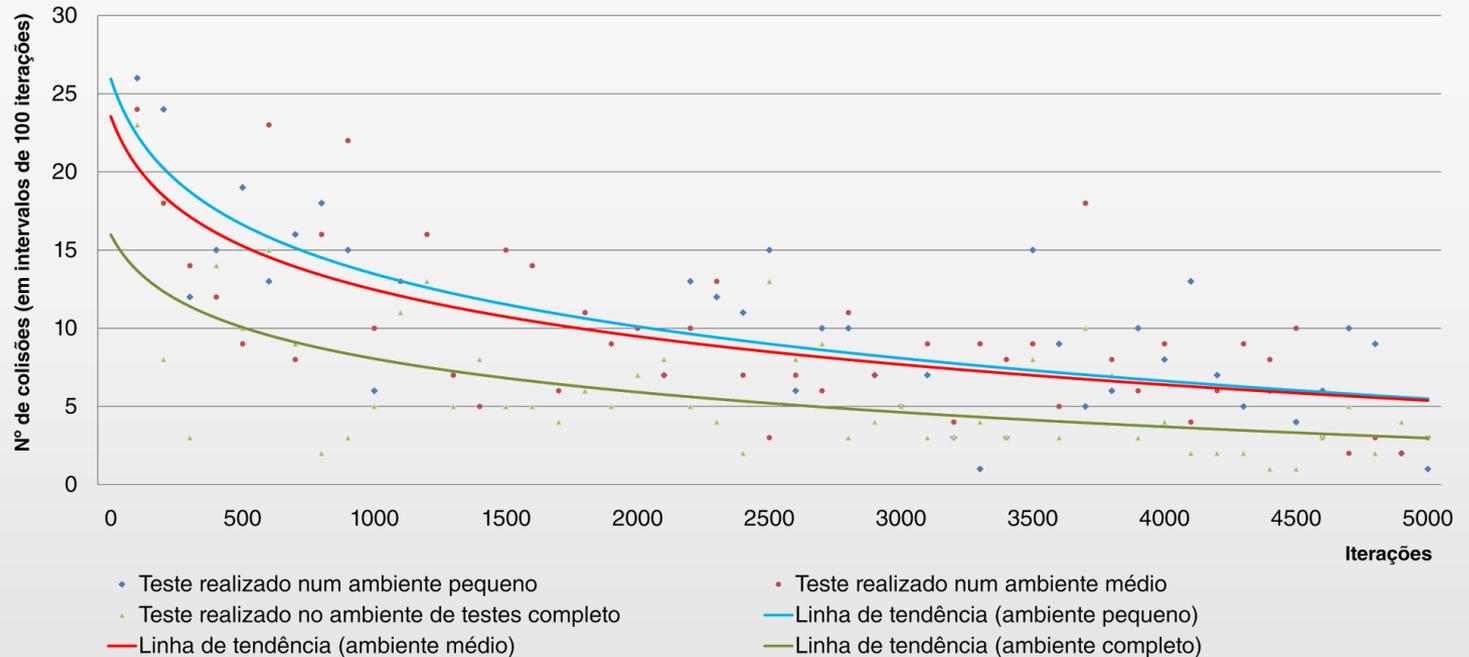
- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros



Resultados da aprendizagem



- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

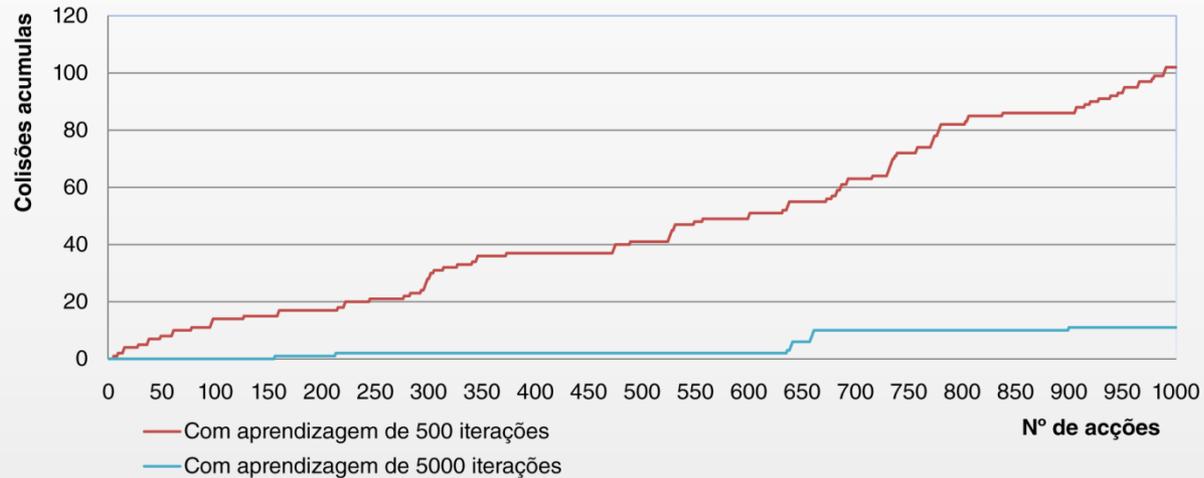


Resultados pós-aprendizagem

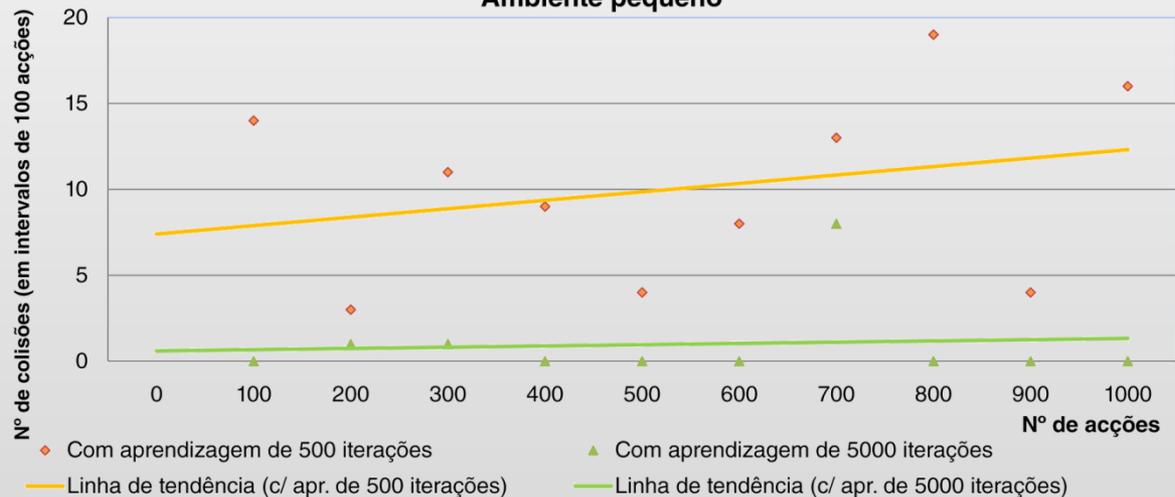


- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

Ambiente pequeno



Ambiente pequeno

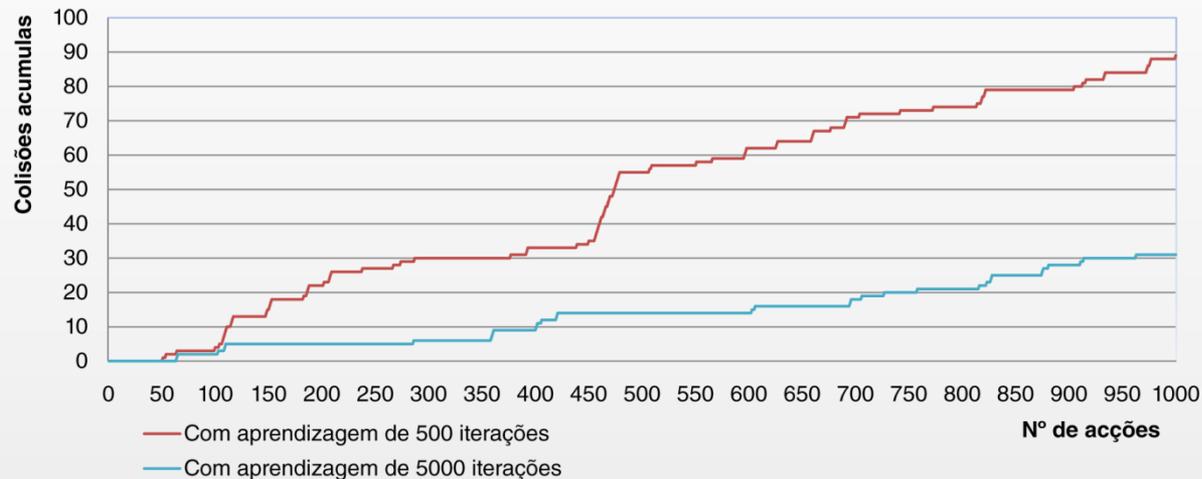


Resultados pós-aprendizagem

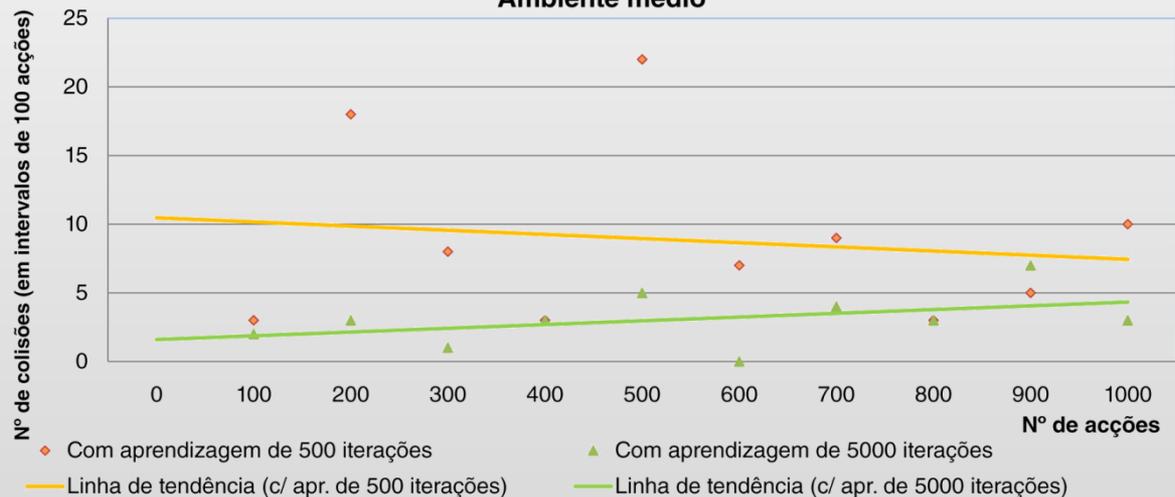


- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

Ambiente médio



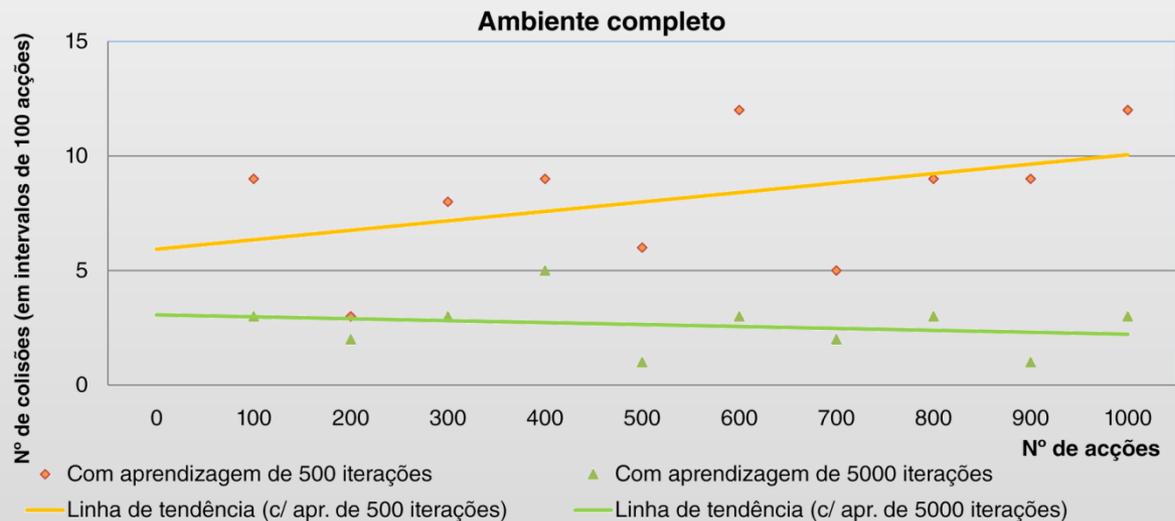
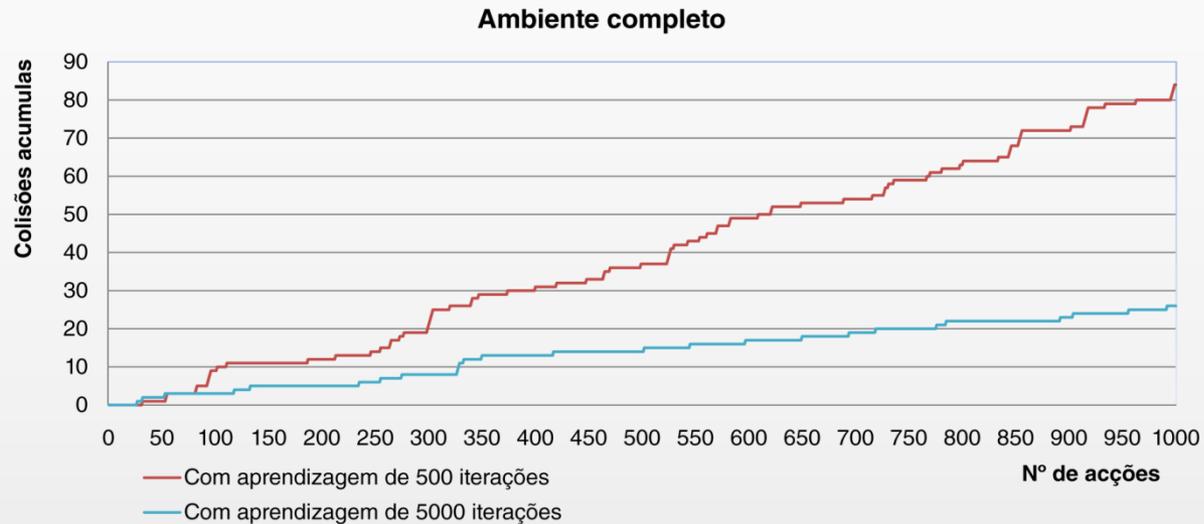
Ambiente médio



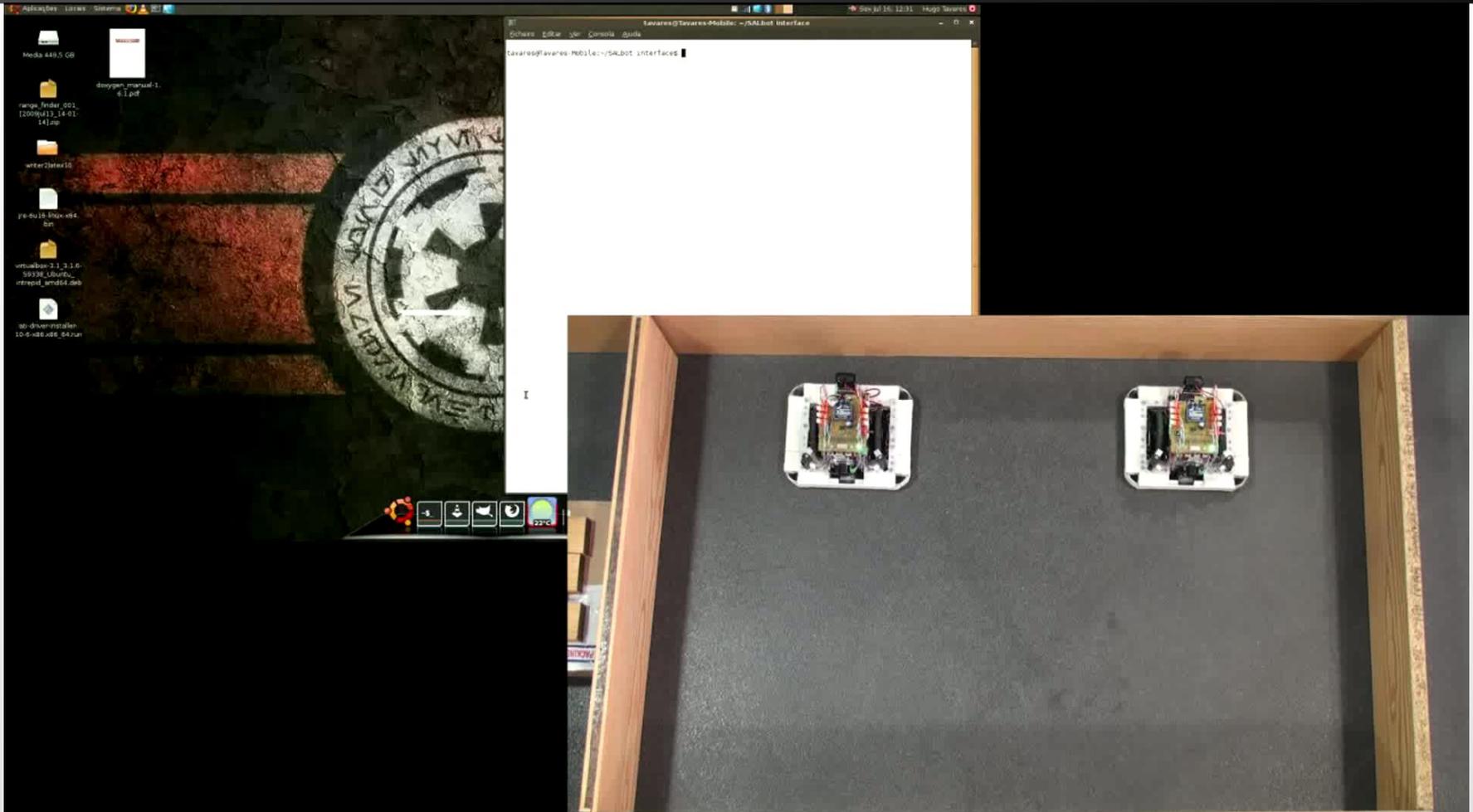
Resultados pós-aprendizagem



- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros



Testes e resultados



Conclusões e Trabalhos futuros



- Conclusões:

- Provou-se a capacidade para aprendizagem dos SALbot's;
- Fiabilidade, e robustez, do chassi e do sistema de pára-choques;
- Limitação em relação à precisão dos servos.

- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros

Conclusões e Trabalhos futuros



- Trabalhos futuros:
 - Comparar diversas configurações do algoritmo de aprendizagem implementado;
 - Testar e comparar novos algoritmos de aprendizagem;
 - Integrar novos dispositivos para algoritmos de aprendizagem mais complexos.

- Objectivos e Motivação
- Projecto dos robôs
- Software
- Testes e Resultados
- Conclusões e Trabalhos futuros



Concepção e desenvolvimento de pequenos robôs para auto-aprendizagem

FIM