

## Resenha

---

### Artigo

Bianchi, R. A. C., Ribeiro, C. H. C., e Costa, A. H. R. (2009). On the relation between ant colony optimization and heuristically accelerated reinforcement learning. *1st International Workshop on Hybrid Control of Autonomous System*, pp. 49–55.

### Comentários

O artigo apresenta uma comparação entre algoritmos das classes *Ant Colony System* (ACS) e *Heuristically Accelerated Reinforcement Learning* (HARL) aplicados no domínio do Problema do Caixeiro Viajante (PCV). Para isso, os autores propõem o algoritmo *Heuristically Accelerated Distributed Q-learning* (HADQL).

Formalmente, o HARL usa uma função heurística para influenciar na escolha de ações do agente de aprendizado. Dessa forma, o HADL é um algoritmo HARL estendido do *Q-learning* Distribuído (DQL), com a adoção de uma função heurística para a seleção de ações.

Os experimentos são realizados com o objetivo de comparar os desempenhos dos algoritmos DQL, ACS e HADQL na resolução de instâncias do PCV. Para isso, é adotada a TSPLIB, repositório de problemas do PCV. Os autores consideram 10 instâncias do PCV, com o número de localidades variando entre 48 e 170: berlin52, kroA100, kroB100, kroC100, kroD100, kroE100, kroA150, ry48, kro124 e ftv170.

Os resultados apontam desempenhos similares entre os algoritmos HADQL e ACS. Assim, reforçando a tese dos autores que o ACS pode ser considerado um caso particular do HADQL. Já o DQL apresentou os piores resultados entre os métodos analisados.

Vale ressaltar que, a única combinação de parâmetros de aprendizado por reforço (AR) adotada nos experimentos é:  $\alpha = 0,1$  (taxa de aprendizado),  $\gamma = 0,3$  (fator de desconto) e  $\epsilon = 0,1$  (política  $\epsilon - greedy$ ). Assim, uma possibilidade para trabalhos futuros é analisar o desempenho dos algoritmos HADQL e DQL de acordo com a sensibilidade dos parâmetros do AR. Nessa linha, já foram realizadas algumas publicações do Grupo de Controle e Modelagem (GCOM).