

Accepted Manuscript

Multi-robot inverse reinforcement learning under occlusion with estimation of state transitions

Kenneth Bogert, Prashant Doshi

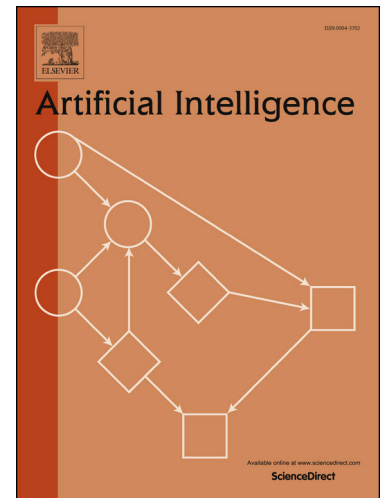
PII: S0004-3702(18)30360-6
DOI: <https://doi.org/10.1016/j.artint.2018.07.002>
Reference: ARTINT 3084

To appear in: *Artificial Intelligence*

Received date: 24 June 2017
Revised date: 27 June 2018
Accepted date: 3 July 2018

Please cite this article in press as: K. Bogert, P. Doshi, Multi-robot inverse reinforcement learning under occlusion with estimation of state transitions, *Artif. Intell.* (2018), <https://doi.org/10.1016/j.artint.2018.07.002>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Multi-Robot Inverse Reinforcement Learning under Occlusion with Estimation of State Transitions

Kenneth Bogert^a, Prashant Doshi^b

^a*Department of Computer Science
University of North Carolina Asheville, Asheville, NC 28804 USA*
^b*THINC Lab, Department of Computer Science
University of Georgia, Athens, GA 30602 USA*

Abstract

Inverse reinforcement learning (IRL), analogously to RL, refers to both the problem and associated methods by which an agent passively observing another agent's actions over time, seeks to learn the latter's reward function. The learning agent is typically called the learner while the observed agent is often an expert in popular applications such as in learning from demonstrations. Some of the assumptions that underlie current IRL methods are impractical for many robotic applications. Specifically, they assume that the learner has full observability of the expert as it performs its task; that the learner has full knowledge of the expert's dynamics; and that there is always only one expert agent in the environment. For example, these assumptions are particularly restrictive in our application scenario where a subject robot is tasked with penetrating a perimeter patrol by two other robots after observing them from a vantage point. In our instance of this problem, the learner can observe at most 10% of the patrol.

We relax these assumptions and systematically generalize a known IRL method, Maximum Entropy IRL, to enable the subject to learn the preferences of the patrolling robots, subsequently their behaviors, and predict their future positions well enough to plan a route to its goal state without being spotted. Challenged by occlusion, multiple interacting robots, and partially known dynamics we demonstrate empirically that the generalization improves significantly

Email addresses: kbogert@unca.edu (Kenneth Bogert), pdoshi@uga.edu (Prashant Doshi)

Download English Version:

<https://daneshyari.com/en/article/6852982>

Download Persian Version:

<https://daneshyari.com/article/6852982>

[Daneshyari.com](https://daneshyari.com)