

Accepted Manuscript

Policy Invariance under Reward Transformations for Multi-Objective Reinforcement Learning

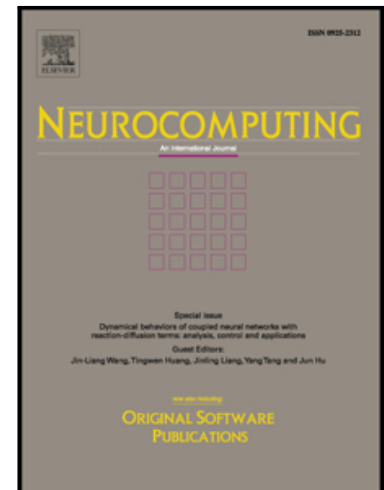
Patrick Mannion, Sam Devlin, Karl Mason, Jim Duggan, Enda Howley

PII: S0925-2312(17)31103-7
DOI: [10.1016/j.neucom.2017.05.090](https://doi.org/10.1016/j.neucom.2017.05.090)
Reference: NEUCOM 18604

To appear in: *Neurocomputing*

Received date: 31 March 2016
Revised date: 13 May 2017
Accepted date: 15 May 2017

Please cite this article as: Patrick Mannion, Sam Devlin, Karl Mason, Jim Duggan, Enda Howley, Policy Invariance under Reward Transformations for Multi-Objective Reinforcement Learning, *Neurocomputing* (2017), doi: [10.1016/j.neucom.2017.05.090](https://doi.org/10.1016/j.neucom.2017.05.090)



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.

Policy Invariance under Reward Transformations for Multi-Objective Reinforcement Learning

Patrick Mannion^{a,b,*}, Sam Devlin^c, Karl Mason^b, Jim Duggan^b, Enda Howley^b

^a*Department of Computer Science & Applied Physics, Galway-Mayo Institute of Technology, Ireland*

^b*Discipline of Information Technology, National University of Ireland Galway, Ireland*

^c*Department of Computer Science, University of York, United Kingdom*

Abstract

Reinforcement Learning (RL) is a powerful and well-studied Machine Learning paradigm, where an agent learns to improve its performance in an environment by maximising a reward signal. In multi-objective Reinforcement Learning (MORL) the reward signal is a vector, where each component represents the performance on a different objective. Reward shaping is a well-established family of techniques that have been successfully used to improve the performance and learning speed of RL agents in single-objective problems. The basic premise of reward shaping is to add an additional shaping reward to the reward naturally received from the environment, to incorporate domain knowledge and guide an agent's exploration. Potential-Based Reward Shaping (PBRS) is a specific form of reward shaping that offers additional guarantees. In this paper, we extend the theoretical guarantees of PBRS to MORL problems. Specifically, we provide theoretical proof that PBRS does not alter the true Pareto front in both single- and multi-agent MORL. We also contribute the first published empirical studies of the effect of PBRS in single- and multi-agent MORL problems.

Keywords: Reinforcement Learning, Multi-Objective, Potential-Based, Reward Shaping, Multi-Agent Systems

*Corresponding author

Email addresses: patrick.mannion@gmit.ie (Patrick Mannion), sam.devlin@york.ac.uk (Sam Devlin), k.mason2@nuigalway.ie (Karl Mason), jim.duggan@nuigalway.ie (Jim Duggan), ehowley@nuigalway.ie (Enda Howley)

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات