Multi-Objective Multiagent Credit Assignment in NSGA-II Using Difference Evaluations

(Extended Abstract)

Logan Yliniemi
Oregon State University
Corvallis, OR, USA
logan.yliniemi@engr.orst.edu

Drew Wilson
Austin Peay State University
Clarksville, TN, USA
wilson.andrew008@gmail.com

Kagan Tumer
Oregon State University
Corvallis, OR, USA
kagan.tumer@oregonstate.edu

ABSTRACT

Determining the contribution of an agent to a system-level objective function (credit assignment) is a key area of research in cooperative multiagent systems. Multi-objective optimization is a growing area of research, though mostly focused on single agent settings. Many real-world problems are multiagent *and* multi-objective, (e.g., air traffic management, scheduling observations across multiple exploration robots) yet there is little work on their intersection.

In this work, we leverage recent advances in single-objective multiagent learning to address multi-objective domains. We focus on the impact of difference evaluation functions (which extracts an agent's contribution to the team objective) on the Non-dominated Sorting Genetic Algorithm-II (NSGA-II), a state-of-the-art multi-objective evolutionary algorithm. We derive multiple methods for incorporating difference evaluations into the NSGA-II framework, and test each in a multiagent rover exploration domain, which is a good surrogate for a wide variety of distributed scheduling and resource gathering problems. We show that how and where difference evaluations are incorporated in the NSGA-II algorithm is critical, and can either provide significant benefits or destroy system performance, depending on how it is used. Median performance of the correctly used difference evaluations dominates best-case performance of NSGA-II in a multiagent multi-objective problem.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent Systems

Keywords

Multiagent Learning; Multi-Objective Problem; Multi-Objective Evolutionary Algorithm; NSGA-II

1. INTRODUCTION

Cooperative multiagent systems focus on determining how best to employ all agents in a system to efficiently produce a desirable system level outcome. A key step in this process is the *credit assignment problem*, where the contribution of each agent to the system is assessed. Credit assignment operators have been studied in a wide variety of experimental domains [1]. However, in each of these cases, the agent optimize a single well-defined objective function. In the real world, it is unlikely that a single value can be optimized while ignoring all other concerns, and instead, multiple must be considered simultaneously. In non-multiagent problems, algorithms have been developed for handling multiple objectives simultaneously. One of the most successful of these is the Non-dominated Sorting Genetic Algorithm-II (NSGA-II) [2], which has been used in a very wide variety of applications, from facial recognition to HIV therapy to rain water reuse.

Many interesting problems involve multiple agents and multiple objectives. In this work we focus on the impact of difference evaluations — a state-of-the-art credit assignment operator — on NSGA-II. We show that difference evaluations provide complimentary benefits to NSGA-II in a multi-objective multiagent system.

2. DIFFERENCE EVALUATION FUNCTIONS

The **global evaluation function** (G) is the system performance of the team as a whole. Training on this signal encourages the agent to act in the system's interest, but includes a large amount of noise from other agents acting simultaneously.

The **difference evaluation function** (D_i) is a shaped reward signal that helps an agent quickly learn the consequences of its actions on the system [1]. It is defined as:

$$D_i(z) = G(z) - G(z_{-i}) \tag{1}$$

where $G(z)$ is the global system performance for the system considering the joint state-action z, and $G(z_{-i})$ is $G(z)$ for a theoretical system without the contribution of agent i. Any action taken to increase D_i simultaneously increases G, while agent i's impact on its own reward is much higher than its relative impact on G [1].

3. NSGA-II

NSGA-II functions on a two-stage sorting operator. For each point it calculates a "non-domination rank" based on the points which dominate it, and a "crowding distance", based on its proximity to other points with the same non-domination rank. Points are sorted first by non-domination rank, and secondarily by crowding distance [2]. Difference evaluations require a real value to function as defined, so we first derived a real-valued function that provides an equivalent total order of policies to NSGA-II.

4. EXPERIMENTAL RESULTS

We present Empirical Attainment Functions [3] over 100 statistical runs for NSGA-II in a multiagent system using (Fig. 1) Global Evaluations, (Fig. 2) Difference Evaluations before NSGA-II, and (Fig. 3) Difference Evaluations after NSGA-II. The team's goal is to maximize the each objective of the two objectives, so an EAF that covers more area in the objective space is better performance.

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

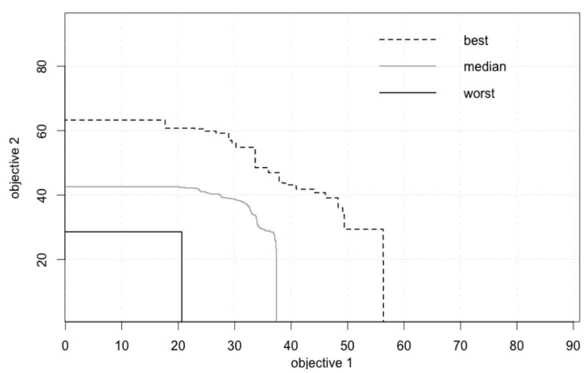

Figure 1: Global NSGA-II EAF; Decentralized.

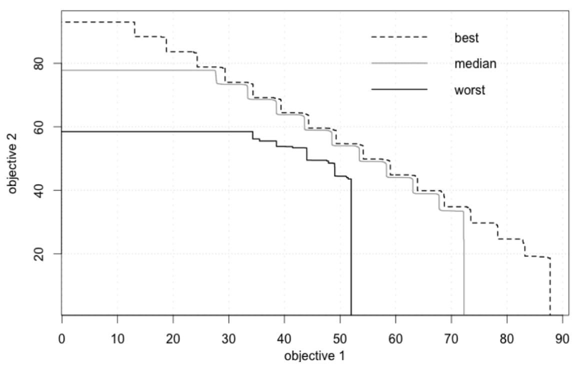

Figure 2: NSGA-II Calculation after Difference Evaluation EAF; Decentralized

The domain is a version of the Continuous Rover Domain [1] with two different types of data that the team must collect simultaneously. We simulated a team of 10 rovers observing 50 POIs that each contain one of two types of data.

Global Evaluation.

Figure 1 shows that the team achieves moderate performance using the global evaluation, but due to the credit assignment problem, the individuals on the team cannot effectively determine their best policies, and the team's performance as a whole suffers.

Difference → NSGA-II.

Figure 2 shows that the team achieves significantly better performance by first calculating the difference evaluation on each objective individually, and then using these values in an NSGA-II calculation. This effectively solves the credit assignment problem, and creates worst-case performance that dominates best-case performance using the global evaluation.

NSGA-II → Difference.

Finally, Figure 3 shows that if the order of operations of NSGA-II and difference evaluations are simply reversed, taking a "difference of NSGA-II values", produces catastrophic effects on the overall system performance. This is due to the underlying structure of NSGA-II, where the value of a particular solution is not a function of its own values so much as it is a function of the points that dominate it, and its neighbors in the objective space. This creates a series of perfectly flat plateaus, and each agent individually does not have enough effect on the system performance to move the performance to a different plateau. This results in equivalent (zero) feedback given to many agents, even those who are improving system performance on both objectives simultaneously.

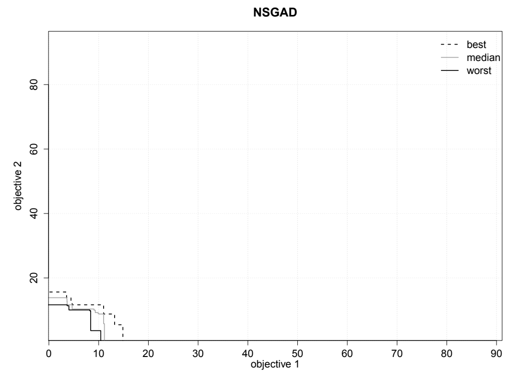

Figure 3: NSGA-II Calculation before Difference Evaluation EAF; Decentralized

5. CONCLUSION

In this work we have presented a novel method for integrating the successful multi-objective algorithm NSGA-II into multiagent systems. We first derived a real-valued fitness assignment evaluation that is equivalent to NSGA-II for use with additional calculations. We then used this formulation in tandem with difference evaluations, and showed that difference evaluations and NSGA-II can provide complimentary benefits to system performance: difference evaluations address the credit assignment problem, while NSGA-II effectively handles multiple objectives simultaneously. We also discovered that because of the formulation of NSGA-II, where many neighboring points in the objective space are valued equivalently, the order of operations when incorporated with difference evaluations is paramount, and can either lead to strong benefits to system performance, or destroy system performance.

We show that credit assignment is of paramount importance in a multiagent, multi-objective setting, and mechanisms must be used that address both problems simultaneously. Furthermore, the mechanisms used to address these problems may have unforeseen interactions.

We are currently expanding this work to consider other multi-objective evolutionary algorithms and their interactions with credit assignment and fitness shaping.

6. ACKNOWLEDGEMENTS

This work was partially supported by the National Energy Technology Laboratory, grant no. DE-FE0012302; and the National Aeronautics and Space Administration, grant no. NNX14AI10G.

REFERENCES

[1] A. K. Agogino and K. Tumer. Analyzing and visualizing multiagent rewards in dynamic and stochastic environments. *JAAMAS*, 17(2):320–338, 2008.

[2] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast elitist multi-objective genetic algorithm: NSGA-II. *Evolutionary Computation*, 6:182–197, 2002.

[3] C. M. Fonseca, A. P. Guerreiro, M. Lopez-Ibanez, and L. Paquete. On the computation of the empirical attainment function. *LNCS*, 6576:121–135, 2011.