Outlook: Using Awareness to Promote Richer, More Human-Like Behaviors in Artificial Agents

Awarded: Most Visionary Paper, 2016

Logan Yliniemi
University of Nevada, Reno
logan@unr.edu

Kagan Tumer
Oregon State University
kagan.tumer@oregonstate.edu

ABSTRACT

The agents community has produced a wide variety of compelling solutions for many real-world problems, and yet there is still a significant disconnect between the behaviors that an agent can learn and those that exemplify the rich behaviors exhibited by humans. This problem exists both with agents interacting solely with an environment, as well as agents interacting with other agents. The solutions created to date are typically good at solving a single, well-defined problem with a particular objective, but lack in generalizability.

In this work, we discuss the possibility of using an awareness framework, coupled with the optimization of multiple dynamic objectives, in tandem with the cooperation and coordination concerns intrinsic to multiagent systems, to create a richer set of agent behaviors. We propose future directions of research that may lead toward more-human capabilities in general agent behaviors.

1. INTRODUCTION

Agents don't act like humans. To a certain extent, this is a desirable trait. Humans can be seen as irrational, moody, and on occasion downright unpleasant.

The agent-based research community has developed compelling solutions for a wide variety of problems, ranging from systems to catch poachers [50] to robotic soccer [3] to stock trading [4] to air traffic management [46, 51], space exploration [33, 52], and many others. However, the solutions produced by the agent research community don't tend to resemble the human decision making process.

Research in this matter in the artificial intelligence community has existed for many decades, with a number of different forms. Common sense [27, 28, 29], context and awareness [5, 13, 14, 39] and lifelong learning [8, 45], are all different instantiations of this concept, which at its core is trying to capture the incredible flexibility and often (apparent) unpredictability of human decision making.

This is not to say that the human way of thinking is somehow superior to agent-based reasoning, but instead is to ask why we cannot achieve this in addition to the advantages that agents have in solving complex problems.

In this work we posit that there may be two prongs which form a very simple answer: first, that the solutions simply do not exist within the paradigm that we, as a community, have been using to solve these problems; and second, that rich decision making requires a broader sense of awareness of one's environment and its meaning, which has not yet received research attention.

At their most basic, most papers in the field produce some form of *agent* to solve some *problem*. Over the years, we've created more-and-more impressive *agents* to solve increasingly difficult *problems*. This is the tried and true framework for agent-based research. Find a problem, and specifically tailor an agent-based algorithm to solve this problem.

Despite, or possibly because of these successes, the community has not made significant steps toward the richer set of behaviors that humans exhibit on an everyday basis. Perhaps it is not the pursuit of a particularly impressive agent to solve a particularly difficult task which will lead us toward agents which exhibit these rich behaviors we seek, but instead these behaviors may require a paradigm change.

This type of creative barrier is one that is mirrored in another field: optimization. Many optimization techniques have been developed for a wide variety of optimization problems, but when optimizing a single value, there are only so many behaviors that can be described this way, and thereby discovered by a single-objective optimization. In recent years, complex optimization problems are not solved by an excessively impressive *optimizer* solving a difficult *problem*, but instead, through a different paradigm. Multi-objective optimization offers a much richer set of behaviors that describes a more complete set of desirable behaviors a system may exhibit [30, 34].

To a certain extent, this is a leap that the agent-based research community is and has been making. We've discovered that some of our techniques from single-objective problems are applicable to multi-objective problems [53, 56], and even that creating a multi-objective problem can make the single-objective problems easier to solve [6, 7].

However, in this work we argue that the crux of the issue does not lie in considering agent-based problems as multi-objective problems, as this only addresses a portion of the larger issue. We posit that human decision making can be reasonably modeled by a multi-objective process, with constantly-shifting, dynamic, non-linear priorities. We pose a series of human experiences that illustrate this point, and use these experiences to form a paradigm through which each of these issues can be addressed by the agent-based research community.

The remainder of this work is organized as follows: we begin in Section 2 by offering some background on multi-objective optimization, since this is a central tenet of our outlook. We then identify a series of human experiences in Section 3 that support the dynamic multi-objective model of a human. In Section 4 we begin building an agent-based framework that can reflect this process and identify some portions of the work that are being done. Finally, in Section 5, we conclude this work with a challenge to the agent community to reach this vision.

2. BACKGROUND: OPTIMIZATION

Within the context of this work, it is important to understand the beginnings of multi-objective optimization (Section 2.1), its modern presence (Section 2.2), and how the form of the reward can change the behavior (Section 2.3). However, we begin by discussing the general concept of optimization.

The core concept of single objective optimization is to choose a set of parameters which you have control over, $\vec{x}$, such that you can either minimize or maximize a value you can't directly control, y, through some form of functional mapping $y = f(\vec{x})$. $f(\vec{x})$ can be nonlinear, discontinuous, stochastic, and difficult or expensive to sample, which form some of the core issues that has kept the field of optimization vibrant and active for many years.

2.1 History of Multi-Objective Optimization

Though many concepts in the field of multi-objective problem solving are named after Vilfredo Pareto, we traced the origins of the field beyond Pareto, to Edgeworth [11].

Edgeworth establishes that, given the choice between a large quantity of good A and a small quantity of good B, or a small quantity of good A and a large quantity of good B, an individual might be indifferent to which set of goods he receives. This establishes the concept of an indifference curve (a curve along which one combination of goods is not preferred to another combination also located on the curve), and also to the concept of a preference curve, which lies perpendicular to the indifference curve.

Pareto solidified the study of the field. He discusses a concept that he calls *ophelimity*, which can be roughly associated with economic use or utility, which he defines as follows [31, 32]:

> *For an individual, the ophelimity of a certain quantity of a thing, added to another known quantity (it can be equal to zero) which he already possesses, is the pleasure which this quanitity affords him*

Pareto makes a strong case that the goal of an individual is to constantly increase their personal ophelimity as far as is feasible. Combining the works of Edgeworth and Pareto, this involves the individual moving along their personal preference curve, which sits perpendicular from his indifference curve, and may be nonlinear.

2.2 Multi-Objective Optimization

Multi-objective optimization is an extension to the single-objective optimization process, where the formulation instead is to maximize or minimize (or some mixture of the two) a vector of solutions $\vec{y} = f(\vec{x})$. Each individual element of $\vec{y}$ can be optimized simultaneously in the formulation discussed in the previous section, but the primary challenge in multi-objective optimization is the optimization of all of these quantities simultaneously. This leads to an entire set of solutions which form the Pareto optimal set on the border between the feasible and infeasible portions of the objective space. These Pareto optimal solutions describe the optimal tradeoffs between the objectives.

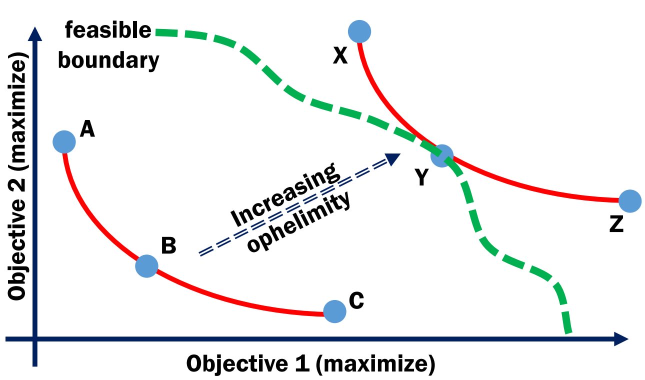

Figure 1: Curve ABC forms an indifference curve, as does XYZ. Curve XYZ represents an increase in ophelimity from ABC. Since Y is the feasible solution with the highest ophelimity, it will be preferred by the decision maker.

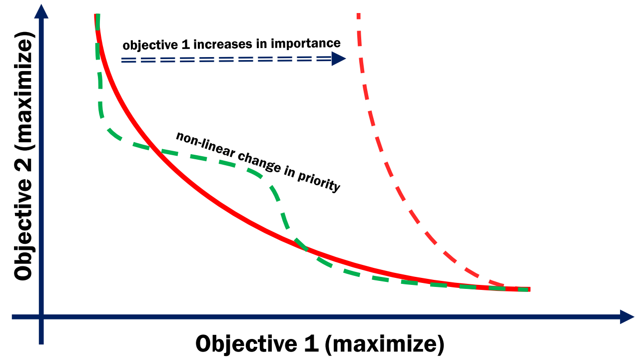

Figure 2: A curve of indifference (solid) can change shape with time. These changes may be easy to parameterize, or nonlinear and difficult to describe, especially with higher numbers of objectives.

This expansion of the problem has led to many advances in optimization, and has allowed solution of extremely complex optimization problems, which would be very difficult to pose in a single-objective sense [16, 21, 49].

Methods in multi-objective optimization vary widely. The simplest and possibly widest-used is the linear combination, in which the (often weighted) objectives are simply added together. This is very computationally efficient, but has well-documented drawbacks, and does not provide the richer behavior space we seek [2, 9, 25]. The linear combination can provide these richer behavior spaces if combined with the concept of indifference, and if the objective space is transformed to guarantee specific types of convexity [54, 55].

Other concepts include nonlinear schemes [18, 26], partitioning the search space [36, 37], and population-based methods in which each population member is compared (pairwise) to each of the other population members, to develop some fitness metric [10, 57].

2.3 Rewards Change Behaviors

Imbuing an adaptive agent with a richer set of possible behaviors poses a difficult problem from the reward design standpoint. While we can typically describe in common language what we would like an agent to do, the act of translating this into a reward or evaluation that leads to this behavior is a difficult process, especially when what you want the agent to do changes over time, or uses on a contextual dependence of events that the agent might not have direct awareness of or the capability to sense.

The design of such a framework, in which agents are able to switch between different contexts and weigh different priorities or objectives with different nonlinear weights, which are simultaneously time-varying, is an extremely difficult design problem with the tools that exist to date.

However, in order to develop this richer set of behaviors that captures the flexibility and emergence that are characteristics of human behaviors, we need to develop techniques for designing such time-varying multiple simultaneous rewards, as well as the algorithms that can use these.

3. SOME HUMAN EXPERIENCES

In this section we pose a series of cases which identify ways in which the use of contextual clues can promote awareness and a shift of mindset in human behaviors. We also present relatively simple cases which are still best described by a combination of multiple objectives.

3.1 Class Begins

Consider a group of students who have shown up a few minutes before a class is due to begin, so they begin interacting with each other about whichever topics are on their mind. There is some signal given, whether by an external cue or by the instructor, that class is about to begin, and the students quiet and begin to listen to the instruction being delivered. If a small group of students continues to speak after class has begun, they may be quieted by their classmates.

This case serves to show that human awareness can lead to swift changes in priorities, and that communication as well as passive observation can lead to a person changing contexts. While this dynamic and the exact mechanics may change on a classroom-to-classroom basis, there is a nearly universally understood "time for outside of class matters" and "time for instruction", each with very different priorities. The shift between these is rapid and shared among the people involved.

3.2 A Loud Noise

Imagine that you are outdoors in a city center, and suddenly, you hear a loud sound. Not only you, but everyone around you, will turn toward the direction of the noise, to determine whether it was a signal of a context switch. In this situation, Shaw states "Unless the danger is very obvious, people often require secondary or confirmatory data before they will begin to react" [40].

Was it simply a car backfiring? Was it a siren? An auto accident? By gathering additional information, you're able to make an intelligent and rational decision about what to do next. Depending on what the additional information shows, your priorities might rapidly shift back to (i) whatever they were previously, especially if there is no perceived change in context; (ii) flight away from the danger; or (iii) to help those in harm's way.

This case serves to show that human awareness detects changes in the environment which signal broader changes in context, and that a change in context can lead to drastically different priorities, which may vary between individuals. It also serves to show that humans use supporting first-hand observations to verify a possible context switch.

3.3 Socially Appropriate Navigation

Consider the simple act of trying to navigate through a crowded hallway in a way that does not disturb those around you. This Socially Aware Navigation is a problem which humans readily solve on a regular basis [15]. In order to properly address this problem, though, you have many competing objectives. As a sample of a set of possible priorities,

i) you are trying to navigate to your goal as quickly as possible
ii) you are trying not to physically disturb any other person along the way
iii) you are trying to avoid walking through groups of people talking with each other
iv) you are trying to expend minimal energy
v) you are trying to stay with your group members
vi) you are preoccupied with your thoughts
vii) you are trying to have courteous interactions

Depending on the details of your situation, your priorities are going to be very different.

- *Efficiency*: If you're having a tough day, perhaps you're much more concerned with (i) and (iv) than the remainders.
- *A hall of coworkers*: If the hall is filled with your colleagues, you may prioritize (vii), along with (iii).
- *Late for an important meeting*: (i) may take precedence over (iii), and you might put no priority at all on (iv).
- *Absentminded*: If other events are occupying your thoughts and attention, you may implicitly place a higher priority on (vi), and allow the others to take lower precedence.
- *A foreigner*: If you are in a foreign place and do not speak the language, you might be more inclined to avoid interactions and therefore prioritize (ii), (iii), and (v).
- *A parent with small children*: (v) likely takes very high priority, with a possible side of (ii) and (iii); you might simply acknowledge that (i) and (iv) are not useful priorities.
- *Inconsiderate others*: If the people crowding the hallway are not being considerate of the people making their way through, perhaps (ii) will take a lower priority in your mind.
- *Combination*: These situations are not mutually exclusive, and if you have a combination of these situations, you may have some combination of the priorities of each.

All of these different sets of priorities are completely rational, though they lead to vastly different courses of action. It is an incredibly human trait that we each can look at the same situation, and, based on our previous experiences and current priorities, come to a different conclusion about the actions that should be taken. This is also why it is so easy to think that someone else is making the wrong choice in a situation. If we are weighing their actions and the likely outcomes with our own priorities, then it is extremely likely that they may appear irrational. They could be using a different prioritization of the same objectives that we are considering, but it is possibly more likely that they are trying to optimize an objective that we haven't even considered in the first place.

To compound this problem, interacting within the human environment is an extremely information-limited problem. It is difficult, even with prolonged shared experiences, to completely understand the motivations and past experiences of those around us, which inherently guide their priorities within a situation. Finally, very different mindsets can lead to the same behaviors: an absentminded person could behave similarly to one concerned only with their path efficiency. They have very different motivations, and different priorities as expressed above, but could exhibit similar observable behaviors.

This case serves to show that with different sets of priorities, different action sets can be seen as equally rational and reasonable. Additionally, without thoroughly understanding an individual's priorities, judging the rationality of their actions is extremely difficult.

3.4 Falsely Shouting "Fire" in a Theatre

Consider, for a moment, the concept of a person entering into a crowded theatre and shouting "Fire!" when there is none. For a moment, the theatre goers may briefly be confused, as the exclamation does not fit into the context that they were expecting. Is this a part of the play? Then, after a short time to process, each individual may rapidly change their priorities, from maximizing their enjoyment to minimizing their time inside the theatre. This process can happen rapidly in parallel, creating a mass panic.

In a decision from 1919, the U.S. Supreme court noted that this is one of the (very few) exceptions to free speech under the U.S. constitution. To quote the decision: "The most stringent protection of free speech would not protect a man in falsely shouting fire in a theatre and causing a panic. It does not even protect a man from an injunction against uttering words that may have all the effect of force" [20].

This decision cites that the use of words may have all the effect of force, and the reason for this is the rapid and extreme context switching that would happen for each person sitting in the theatre. It immediately places every person in the theatre in danger from the circumstances that may arise from the mass exodus from the theatre by (reasonably) self-concerned patrons. In fact, simply shouting fire has led to a loss of life in the panic of some situations [35, 48], whereas in other highly dangerous situations that actually involved a large fire, no loss of life occurred [40].

This case serves to show that a human's sense of context can be manipulated by the actions of others, and that the sense of context has a high impact on the actions of others: "all the effect of force".

4. TOWARD RICHER AGENT BEHAVIORS

In order to achieve rich behaviors such as these, a possible route is to create a framework which has the same characteristics, both within sensing the context and when it changes, and in the decision making process once a context has been identified.

These characteristics are:

Context sensing

- Independent detection of a context change
- Inter-agent communication to facilitate context switching
- Sensory verification of a communicated context switch

Decision-making

- Event-dependent multiple priorities
- Priorities with nonlinear preference curves
- Varying priorities based on past experiences
- A strong change in behavior corresponding with changes in context

In this section, we identify areas in which the MAS and AI communities have made some steps toward imbuing agents with these characteristics, and some possible future directions of research. This is linked to over 3 decades of work [23, 24] in awareness, long term autonomy, and common sense for artificial intelligence, but in this section we look at the research with an eye toward using multi-objective optimization with dynamically-changing priorities.

4.1 The Detection of Context Changes

Giving an agent awareness of context, which is broader than a simple state representation, is an extremely large research problem. It is possible that contributions to such a detection method this could come from sources like transfer learning [38, 43, 44] anomaly detection [1, 17, 22], the detection of opponent policy switching in non-stationary problems [12, 19, 47] or shared autonomy [41, 42].

Each of these problem types are ones in which the MAS and AI community have many collective years of experience solving. In the particular application of identifying context changes, we propose one avenue: since many candidate priorities must exist for the richer behavior space that we seek, why not constantly track the evaluations of these objectives, and use the past history as a litmus test? If an agent takes an action and can predict a vector of rewards, but receives a vastly different vector, it is very possible that a context change has happened.

4.2 The Use of Context

Once a shift in context has been detected, the agent can suddenly find itself in a world of uncertainty, and there are many research questions to be addressed: how does the agent select its new set of objectives from among the entire set it may consider? How does the agent prioritize these objectives, and with what form of a preference scheme? How can policy information be maintained across changes in context, and still used in a constructive manner?

Again, the MAS and AI community has many collective years of solving these types of problems. The selection of a new set of priorities without excessive regret is in many ways similar to handling a new opponent strategy in a competitive game. The preference scheme can be built up based on what can be achieved within the constraints of the new context. Outside knowledge can be incorporated with reward shaping. Policy information can be maintained through transfer.

The incorporation of any combination of these at once is a large research problem, which requires concerted effort on a community-wide, collaborative level. It requires publishing work that requires the knowledge of multiple sub-fields to properly review and understand. It requires a level of risk. However, it also provides a substantial reward: a future *agent* that can not only solve a particularly difficult *problem*, but can use a sense of awareness to situate itself within its environment, such that it can potentially solve many problems despite (or due to) many changes in context along the way.

5. CONCLUSION

In this work we have identified a challenge for the MAS and AI community: the development of agents with a richer set of behaviors, which may be able to mimic the human decision making process. We have identified, through a series of vignettes, some desirable aspects of the human decision making process, and provided a paradigm through which an autonomous agent-based system might be able to mimic these human behaviors, through the incorporation of a sense of *awareness* into the agents. Such agents will be capable of detecting when changes in their environment, their interaction with the environment, or actions of others indicate a change in *context*, and use this to quickly change the set of *priorities* which they consider. These agents will then consider their *priorities* with some form of *non-linear* preference (and indifference), and take actions based on these priorities and preferences. In order to imbue artificial agents with the flexibility and emergence associated with human behaviors, we, as a community, need to develop each of these techniques, with an eye toward integration with each of the others.

REFERENCES

[1] A. Agogino and K. Tumer. Entropy based anomaly detection applied to space shuttle main engines. In *IEEE Aerospace Conference*, 2006.

[2] T. W. Athan and P. Y. Papalambros. A note on weighted criteria methods for compromise solutions in multi-objective optimization. *Engineering Optimization*, 27(155-176), 1996.

[3] S. Barrett and P. Stone. Cooperating with unknown teammates in complex domains: A robot soccer case study of ad hoc teamwork. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, January 2015.

[4] D. Bloembergen, D. Hennes, P. McBurney, and K. Tuyls. Trading in markets with noisy information: An evolutionary analysis. *Connection Science*, 27(3):253–268, 2015.

[5] P. Brézillon. Context in artificial intelligence: I. a survey of the literature. *Computers and artificial intelligence*, 18:321–340, 1999.

[6] T. Brys, A. Harutyunyan, P. Vrancx, M. E. Taylor, D. Kudenko, and A. Nowé. Multi-objectivization of reinforcement learning problems by reward shaping. In *Neural Networks (IJCNN), 2014 International Joint Conference on*, pages 2315–2322. IEEE, 2014.

[7] T. Brys, M. E. Taylor, and A. Nowé. Using ensemble techniques and multi-objectivization to solve reinforcement learning problems. In *ECAI*, pages 981–982, 2014.

[8] A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E. R. Hruschka Jr, and T. M. Mitchell. Toward an architecture for never-ending language learning. In *AAAI*, volume 5, page 3, 2010.

[9] I. Das and J. E. Dennis. A closer look at drawbacks of minimizing weighted sums of objectives for pareto set generation in multicriteria optimization problems. *Structural Optimization*, pages 63–69, 1997.

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

[11] F. Y. Edgeworth. *Mathematical Psychics: An essay on the application of mathematics to moral sciences*. C. Kegan Paul and Company, 1881.

[12] M. Elidrisi, N. Johnson, and M. Gini. Fast learning against adaptive adversarial opponents. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, Valencia, Spain (November 2012)*, 2012.

[13] Thomas Erickson. Some problems with the notion of context-aware computing. *Communications of the ACM*, 45(2):102–104, 2002.

[14] R. Fagin and J. Y. Halpern. Belief, awareness, and limited reasoning. *Artificial intelligence*, 34(1):39–76, 1987.

[15] D. Feil-Seifer. Distance-based computational models for facilitating robot interaction with children. *Journal of Human-Robot Interaction*, 1(1), 2012.

[16] P. Flener, J. Pearson, M. Ågren, C. Garcia-Avello, M. Celiktin, and S. Dissing. Air-traffic complexity resolution in multi-sector planning. *Journal of Air Transport Management*, 13(6):323–328, 2007.

[17] A. Ghosh and S. Sen. *Agent-based distributed intrusion alert system*. Springer, 2004.

[18] I. Giagkiozis and P. J. Fleming. Methods for multi-objective optimization: An analysis. *Information Sciences*, 293:338–350, 2015.

[19] P. Hernandez-Leal, E. Munoz de Cote, and L. E. Sucar. A framework for learning and planning against switching strategies in repeated games. *Connection Science*, 26(2):103–122, 2014.

[20] O. W. Holmes Jr. U.S. supreme court opinion: Schenck v. United States, 1919.

[21] S. Jeyadevi, S. Baskar, C.K. Babulal, and M. W. Iruthayarajan. Solving multiobjective optimal reactive power dispatch using modified NSGA-II. *International Journal of Electrical Power & Energy Systems*, 33(2):219–228, 2011.

[22] B. Kaluža, G. A. Kaminka, and M. Tambe. Detection of suspicious behavior from a sparse set of multiagent interactions. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 955–964. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[23] D. B. Lenat. Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38, 1995.

[24] D. B. Lenat, R. V. Guha, K. Pittman, D. Pratt, and M. Shepherd. Cyc: toward programs with common sense. *Communications of the ACM*, 33(8):30–49, 1990.

[25] R. T. Marler and J. S. Arora. The weighted sum method for multi-objective optimization: new insights. *Structural and Multidisciplinary Optimization*, 2009.

[26] R.T. Marler and J. S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26:369–395, 2004.

[27] J. McCarthy. *Programs with common sense*. Defense Technical Information Center, 1963.

[28] J. McCarthy. Generality in artificial intelligence. *Communications of the ACM*, 30(12):1030–1035, 1987.

[29] J. McCarthy. Artificial intelligence, logic and formalizing common sense. In *Philosophical logic and artificial intelligence*, pages 161–190. Springer, 1989.

[30] A. Messac and P. D. Hattis. Physical programming design optimization for high speed civil transport (HSCT). *Journal of Aircraft*, 33(2):446–44, March 1996.

[31] V. Pareto. *Manuale di Economia Politica*. Piccola Biblioteca Scientifica. Societa Editrice Libraria, 1906.

[32] V. Pareto. *Manual of Political Economy*. MacMillan Press Ltd., 1927.

[33] M. Pěchouček and V. Mařík. Industrial deployment of multi-agent technologies: review and selected case studies. *Autonomous Agents and Multi-Agent Systems*, 17(3):397–431, 2008.

[34] R. Penn, E. Friedler, and A. Ostfeld. Multi-objective evolutionary optimization for greywater reuse in municipal sewer systems. *Water research*, 47(15):5911–592, 2013.

[35] L. A. Powe Jr. Searching for the false shout of fire. *Const. Comment.*, 19:345, 2002.

[36] H. Sato. Inverted PBI in MOEA/D and its impact on the search performance on multi and many-objective optimization. In *Proceedings of the 2014 conference on Genetic and evolutionary computation*, pages 645–652. ACM, 2014.

[37] H. Sato. MOEA/D using constant-distance based neighbors designed for many-objective optimization. In *Evolutionary Computation (CEC), 2015 IEEE Congress on*, pages 2867–2874. IEEE, 2015.

[38] S. Sen, M. Sekaran, and J. Hale. Learning to coordinate without sharing information. In *AAAI*, pages 426–431, 1994.

[39] L. Serafini and P. Bouquet. Comparing formal theories of context in ai. *Artificial intelligence*, 155(1):41–67, 2004.

[40] R. Shaw. Don't panic: behaviour in major incidents. *Disaster Prevention and Management: An International Journal*, 10(1):5–10, 2001.

[41] M. Tambe. Electric elves: What went wrong and why. *AI magazine*, 29(2):23, 2008.

[42] M. Tambe, P. Scerri, and D. V. Pynadath. Adjustable autonomy for the real world. *Journal of Artificial Intelligence Research*, 17(1):171–228, 2002.

[43] M. E. Taylor, G. Kuhlmann, and P. Stone. Autonomous transfer for reinforcement learning. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*, pages 283–290. International

Foundation for Autonomous Agents and Multiagent Systems, 2008.

[44] M. E. Taylor and P. Stone. An introduction to intertask transfer for reinforcement learning. *AI Magazine*, 32(1):15, 2011.

[45] S. Thrun and T. M. Mitchell. *Lifelong robot learning*. Springer, 1995.

[46] K. Tumer and A. Agogino. Distributed agent-based air traffic flow management. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, page 255. ACM, 2007.

[47] Z. Wang, A. Boularias, K. Mülling, and J. Peters. Balancing safety and exploitability in opponent modeling. In *AAAI*, 2011.

[48] M.F. Wendland. The calumet tragedy + death of a city in northern michigan, 1913-1914. *American Heritage*, 37(3):39, 1986.

[49] H. Xu, Z. Zhang, K. Alipour, K. Xue, and X.Z. Gao. Prototypes selection by multi-objective optimal design: application to a reconfigurable robot in sandy terrain. *Industrial Robot: An International Journal*, 38(6):599–613, 2011.

[50] R. Yang, B. Ford, M. Tambe, and A. Lemieux. Adaptive resource allocation for wildlife protection against illegal poachers. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 453–460. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[51] L. Yliniemi, A. K. Agogino, and K. Tumer. Evolutionary agent-based simulation of the introduction of new technologies in air traffic management. *Genetic and Evolutionary Computation Conference (GECCO)*, 2014.

[52] L. Yliniemi, A. K. Agogino, and K. Tumer. Multirobot coordination for space exploration. *AI Magazine*, 4(35):61–74, 2014.

[53] L. Yliniemi and K. Tumer. Multi-objective multiagent credit assignment through difference rewards in reinforcement learning. In *Simulated Evolution and Learning*, pages 407–418. Springer, 2014.

[54] L. Yliniemi and K. Tumer. PaCcET: An objective space transformation to iteratively convexify the pareto front. In *10th International Conference on Simulated Evolution And Learning (SEAL)*, 2014.

[55] L. Yliniemi and K. Tumer. Complete coverage in the multi-objective PaCcET framework. In S. Silva, editor, *Genetic and Evolutionary Computation Conference*, 2015.

[56] L. Yliniemi, D. Wilson, and K. Tumer. Multi-objective multiagent credit assignment in nsga-ii using difference evaluations. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 1635–1636. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[57] E. Zitzler, M. Laumanns, and L. Thiele. SPEA2: Improving the strength pareto evolutionary algorithm. *Computer Engineering*, 3242(103), 2001.