Tradeoffs within Representational Behaviors: Multi-Objective Optimization Strategies for Simulated Military Units

Rob Kewley, Chris Argenta, Chris McGroarty

Simlytics.cloud, LLC, Applied Research Associates, Inc., U.S. Army Combat Capabilities Development Command - Soldier Center (DEVCOM SC) Simulation and Training Technology Center (STTC)
rob@simlytics.cloud, cargenta@ara.com, christopher.j.mcgroarty.civ@army.mil

Abstract

Generating representative behaviors for opposing and adjacent forces is critical for training military units in a simulated environment. Humans are often required in-the-loop because automated systems struggle to provide intelligent and adaptive behaviors that represent appropriate military doctrine and can be explained in the After Action Review. Automated solutions often fail to maintain enough diversity to be replayed in successive iterations of the simulated mission without becoming easily predictable. This paper describes the decision-making approach for automated hierarchical planning agents designed to overcome these challenges. Our Mission Command Agents (MCAs) combine constraint optimization for multi-agent goal reasoning and state-space planning for route planning to generate Courses of Action (COAs) for simulated units. We detail extensions in both steps to produce a set of COAs on the Pareto front for multiple practical (e.g., distance), doctrinal (e.g., cover and concealment), and behavioral (e.g., risk-tolerance or aggressiveness) objectives.

We compare COAs generated using this new approach to results of our previous MCA that uses a more traditional approach of aggregating objectives or preferences into a single weighted utility function and optimizing with respect to that joint utility to find multiple optimal and near-optimal solutions. In contrast to previously emitted behaviors, which often reflected many small differences in utility or positioning, analyzing the trade-offs among multiple objectives produces more meaningfully diverse solutions. In an experiential training context, this diversity supports replaying scenarios with automated units that have distinct preferences and priorities. We summarize our results with evaluation metrics showing a set of COAs can appropriately represent doctrinal constraints with diverse, reasonable, and explainable trade-offs between objectives.

Intelligent Course of Action Planning for Military Units

Modern combat operations are reshaping the experiences of small tactical units. Long range surveillance by drones and electronic surveillance, coupled with long range precision fires, forces them to operate distributed – isolated from their larger organizations. To achieve effects, they must be able to coordinate across large distances and rapidly consolidate at decisive locations, quickly dispersing again before the enemy's fires take effect (Scales, 2019). The Institute of Land Warfare takes a further step, calling for our small units to achieve overmatch through realistic training designed to improve leader and team agility and performance (Roper, 2018). Some of this has been codified in the Army's capstone doctrine, calling for the ability to make contact with the smallest unit possible and converge to achieve effects (HQDA, 2022).

Simulation-based training offers an opportunity to accelerate the development of effective small units and leaders; but, the challenges of modern operations strain the current capabilities. In particular, the simulation must present realism with behaviors representative of those expected on the battlefield. This requires Artificial Intelligence- (AI-) driven forces to behave doctrinally during the simulated mission, and to react in real-time to stimulate rapid decision making and execution by these small units (Owens et. al, 2020). However, AI systems are not yet ready to meet this challenge. Current simulations employ scripted behaviors or finite state machines that can be cumbersome and difficult to develop (McGroarty and Gallant, 2022). Machine learning offers promise in this domain (Narayanan et. al. 2021), but the simulation's inability to provide a realistic training environment for the learning systems limits the realism offered by these approaches. In addition, machine learning systems have not been developed to adhere to a specific force's doctrine or to explain their behaviors in the After Action Review.

This paper stems from an effort to develop AI capabilities that overcome these limitations. Mission Command Agents (see Figure 1) are designed based on the Army's Mission Command doctrine for command and control of Army forces (HQDA 2019). They receive a mission order from the training system and begin building situational understanding

[1] Copyright © 2023, by the authors. All rights reserved.

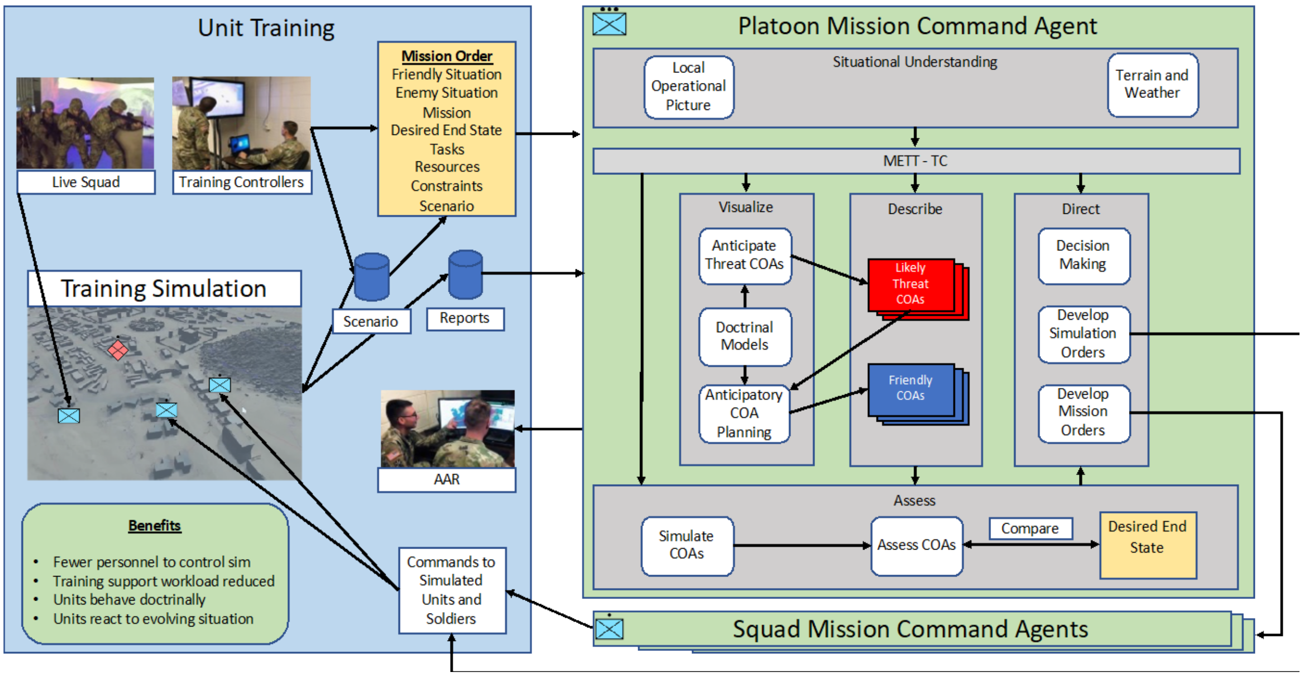


Figure 1. Mission Command Agents for intelligent COA planning for simulated units.

of the current mission. They then visualize how the mission might play out by anticipating threat COAs to consider during the development of friendly COAs. An assessment process provides data to support a decision about which COA to send to the training system for execution (Argenta et. al., 2022). This paper focuses on improving the overall diversity and quality of COAs available to the decision maker (human or automated). The development of Pareto sets during intelligent planning replicates different decision makers with the ability to trade-off between doctrinal objectives. In addition, diversity is encouraged so trainees cannot easily predict the automated forces' actions based on previous simulations.

Generating Representative and Diverse Courses of Action

Given the same set of doctrinal rules, humans may interpret, prioritize, and apply the rules, uncertainties, and personal tolerances to produce different behaviors for the same situation. Decision-making behaviors can be both representative of the rules and diverse in their application. In automated systems, we optimize behaviors with multiple objectives, which may involve trade-offs that allow solutions that results in similarly diverse but representative behaviors.

To intentionally generate behaviors that reflect the ranges a human might select, it helps to understand which trade-offs exist and how diverse interpretations, prioritizations, and applications of rules might be considered beneficial.

In our problem space, an MCA makes two choices. First, where to position a unit such that the combination of units produces the desired and representative effect against an uncertain picture of what they anticipate the opposing force will do. Second, what route the unit will take to get to the desired position from its current position, which must also represent how that action might be expected to be done by the specific type of unit being automated.

Comparing a Priori to a Posteriori Strategies

The initial implementation of the MCA developed a set of unit position plans for a military offensive operation using a Tabu Search and a custom Best-First State-Space Search for router planning. Both these algorithms transform a multi-objective problem into a single-objective problem by applying *a priori* knowledge of preferences as a set of weights. Our MCA optimized nine different objectives based on relative weights manually selected for that situation. Additionally, these approaches were customized to find multiple optimal and near optimal solutions based on these objectives.

Here, we compare that initial MCA with a new design that uses a second class of techniques that involve finding the Pareto Front and exploring tradeoffs. These *a posteriori techniques* have several advantages. They do not require preferences in advance, they can use population-based techniques to evaluate many solutions in a single iteration, and they offer a meaningfully distinct set of choices and trade-offs to the decision maker.

Optimizing Tactical Position Planning

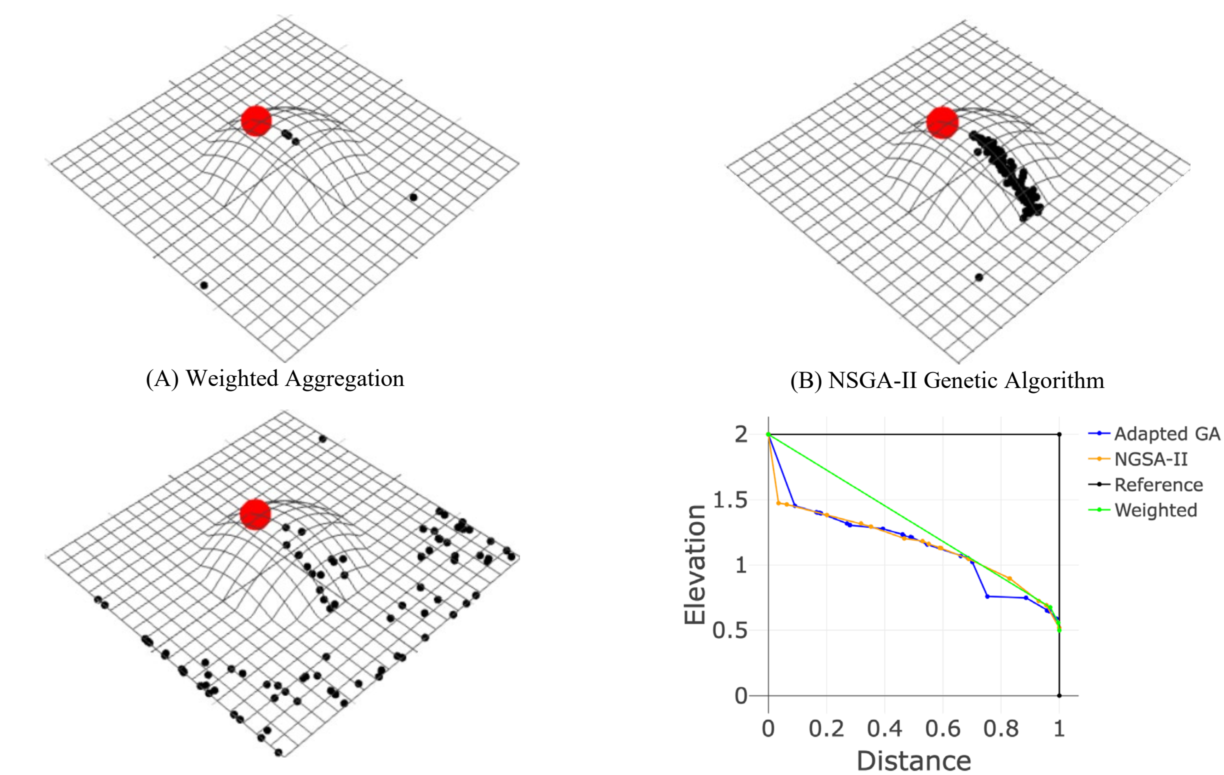
A simplified positioning problem is shown in Figure *2*. Higher elevations allow better observation of the enemy, but our unit must adapt to an already established friendly position (red dot). It is also important for positions to be dispersed. However, once the positions are at least two units apart, further dispersion provides no value. The objectives for the position planner are to maximize elevation and maximize distance from the red dot. Also, note the large flat area for which all elevation values are equal.

The positions in the Weighted Aggregation plot (A) in Figure 2 represent multiple runs using different a priori preferences expressed as weights. This technique allows some exploration of the Pareto frontier and solution space but requires guessing appropriate weights without knowledge of how the objectives interact.

A common algorithm for finding the Pareto Front is the Nondominated Sorting Genetic Algorithm – II (NGSA-II) (Deb et. al., 2002). This algorithm favors nondominated solutions along with solutions that are not near other solutions on the Pareto front. The NGSA-II plot (B) in Figure *2* results from running a population of 100 solutions for 2500 generations. Note that the Pareto Front along the $x = 0$, $0 < y < 1$ line is well defined while the entire flat space where the separation distance $d > 2$ has equal Pareto optimality, but only one point represents this space. It has diversity in the Pareto (objectives) space, but not in the domain (XY) space. In tactical planning for military units, diversity of positioning on the battlefield is at least as important as diversity of preferences.

In order to remedy this situation, we adapted the NSGS-II algorithm to incorporate k-nearest-neighbor crowding distances in both the objective space and XY space. This favors nondominated solutions that are not near other solutions on the Pareto front and not near other solutions in XY space. The results are shown in the adapted genetic algorithm plot (C) of Figure *2* and compared in (D). Note the dispersion along the Pareto front and in XY space, which would appear meaningfully different to military planners.

The performance metrics in Table 1 confirm what we can see visually. The normalized hypervolume metric captures how closely the resulting Pareto front pushes to the "ideal" reference front where maximum values are achieved for each objective (Nebro et. al., 2009). Pareto fronts for each metric are plotted in Figure *2*. The normalized hypervolume value is the percentage of the area under each front, divided by the area under the reference front. For this metric, the weighted aggregation method performs best because its scores lie directly on the Pareto front, covering most of its volume. Note that the diversity seeking elements of the other algorithms yield solutions that are not necessarily exactly on that front.



(A) Weighted Aggregation

(B) NSGA-II Genetic Algorithm

(C) Genetic Algorithm Adapted for Spread in XY Space

(D) Pareto Fronts

Figure 2. Simplified positioning problem where positions (small black dots) seek to maximize elevation and distance from red dot.

The next metric is a spread metric measured in objective distance along the Pareto front. This metric captures the degree to which solutions along the Pareto front are evenly spaced from each other and from the extreme values of that front. Lower scores are better. For this metric, the NGSA-II algorithm performs best, as expected. A similar spread metric captures the degree to which solutions are equally spaced in the XY dimension. The diversity we see for the adapted genetic algorithm is captured by this metric.

	Weighted	*NGSA-II GA*	*Adapted GA*
Norm. Hypervolume	**0.76**	0.59	0.61
Spread Pareto (↧)	1.11	**0.73**	0.84
Spread XY (↧)	0.25	0.33	**0.11**

Table 1. Pareto optimization performance measures for the simplified single unit position problem
↧ - lower is better

A visual inspection of these results yields additional important insights. We expect our Mission Command Agents to behave as real units would in the field. In this case, they will find good solutions, but not necessarily the absolute best. In turn, we accept being a little bit off the Pareto front because it yields diversity of choice and diversity from iteration to iteration. In this sense, the adapted genetic algorithm plot looks more realistic than the others.

To demonstrate these strategies on more complex military planning problems, we placed four units in this same space while trying to simultaneously maximize the mean elevation of all units and the distance between the two closest units. Table 2 shows the performance measures for this problem. The relative performance of the Adapted GA is similar to its performance for the simple problem. It does the best in diversity in the XY space, sacrificing a bit in normalized hypervolume and Pareto spread. For the military planning domain, this is an acceptable tradeoff that allows diversity of choice and solutions that look different to a military trainee.

	Weighted	*NGSA-II GA*	*Adapted GA*
Norm. Hypervolume	**0.72**	0.54	0.61
Spread Pareto (↧)	0.85	**0.78**	1.16
Spread XY (↧)	0.19	0.1	**0.04**

Table 2. Pareto optimization performance measures for placement of four units.

Extending the Position Decision Problem

In our military tactical example, we positioned 4 units with different weapons capabilities (SBF - Support By Fire and ASSLT - Assault) within a given scenario terrain (Figure *4*). A position for an opposing force was given such that the value of various positions was dependent on their relationship to this opposing force as well as selected friendly positions. In this optimization, we used six objectives but simplified to three for presentation purposes, which include:

1. Maximize the cover and concealment – for all friendly positions with respect to the opposing force.
2. Maximize the suppression – the potential firepower that the units can bring to bear on the opposing force position given their weapon characterizations.
3. Maximize the separation – this reflects the distance between units with diminishing returns.

We generated 50 potential position combinations near the Pareto Front with respect to these objective dimensions, as well as the equivalent of the weighted score. A comparison of these solutions is plotted as Figure 3. It is immediately obvious that there is not a single positioning solution that dominates the others in all objectives.

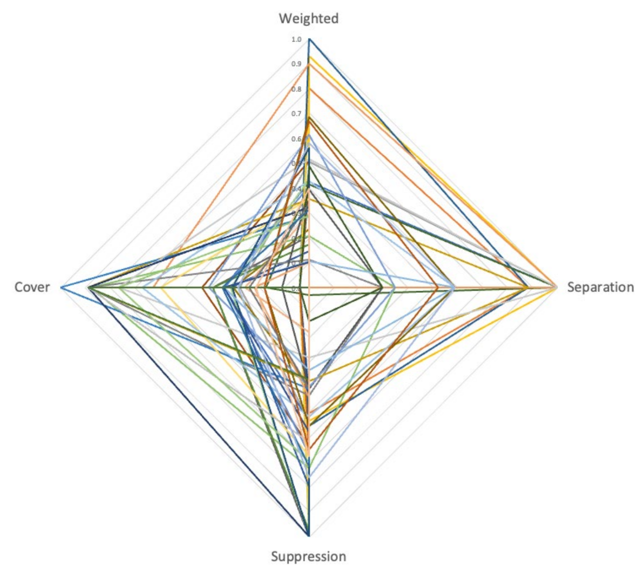


Figure 3. Radar chart showing scores (*larger is better*) of Top-50 solutions for planning 4 units with three objectives. Also shows the trade-off in separation vs. suppression potential.

We quantify the trade-offs between objectives by computing the correlation of scores between each objective (Table 3). Solutions with high total weights tended to trade-off cover as shown by the negative correlation.

	Weighted	*Separation*	*Suppression*	*Cover*
Weighted	1.000			
Separation	0.415	1.000		
Suppression	0.092	-0.459	1.000	
Cover	-0.514	-0.015	-0.070	1.000

Table 3. Correlation matrix for objectives in the four-unit positioning problem. There is a weak negative correlation between separation and suppression. The weighted objective correlates positively with suppression and negatively with cover.

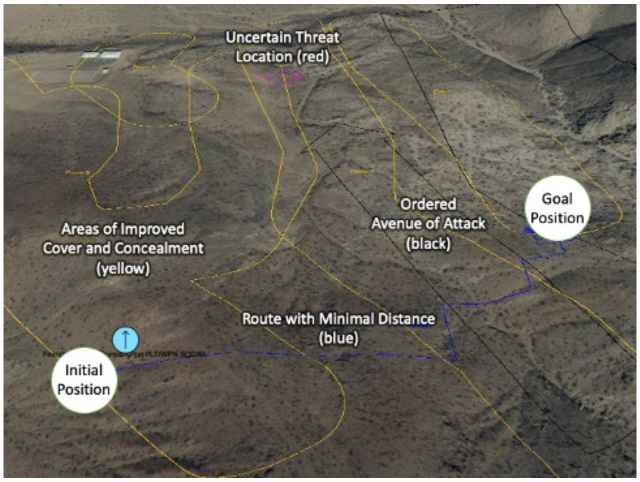


Figure 4. Route planning problem has initial (left white) and goal positions (right white), and a potential position of a threat. Terrain features of interest include draws and crests (yellow), which are doctrinally identified areas for good cover and concealment, and a tactical graphic (black) indicating the avenue of approach this unit was assigned in its mission orders. An example route (blue) is optimal for minimizing distance traveled.

Optimizing Tactical Routes

Route planning is a well-studied problem often employing dynamic programming for a state-space search using an action model (summarized in Russel 1995). In an efficient approach, such as A* (Hart, 1968), there is a single objective cost function against which potential solutions are searched, and an admissible heuristic for estimating the cost remaining until a goal is reached. Much like the positioning problem, combining multiple objectives is often accomplished with an a priori weighting and it can be difficult to predict how objectives will compete for any specific problem instance. There may also be many, potentially different, solutions that score equally optimal with respect to the weighted utility.

In addition, it is difficult to fix weights that consistently produce behaviors that reasonably represent those of an expert because the expert's optimal choice may be adjusted based on the results that they might expect given alternative priorities. This circularity occurs because the expert considers trade-offs based on the expected results (e.g., mentally simulating each COA with various priorities or risks). To make this challenge worse, an expert would likely jointly consider positional options along with route planning.

In the following subsections, we explore how to discover trade-offs in the multi-objective route planning problem by comparing collections of optimal and near-optimal solutions. By comparing a diversity of good (with respect to various strategies) route solutions generated, we discover the trade-offs between the objectives that can be quantified by analyzing the correlation of their effects on the solutions.

Route Planning Decision Problem

We use a simplified route planning problem is shown in Figure *4* on which the results included here are based. In this problem, there is one agent with the ability to move between grid squares (four-way connectivity) with the goal to move from an initial given position to a goal position provided.

There are three objectives being optimized:

1. Minimize the <u>distance</u> traveled – each move action adds a penalty based on the distance between grid centers.
2. Maximize the <u>cover</u> and concealment – amount of time spent in areas identified as crests or draws (doctrine).
3. Maximize the consistency to the mission <u>order</u> – time spent in our prescribed avenue of approach.

The minimum distance route between the initial and goal positions is 15 moves but includes movements through areas that are not concealed or near the ordered avenue of approach. There are over 1,000 unique routes that are equally optimal with respect to only minimizing distance.

Routes with Equal Weighting of Objectives

We have developed a state-space route planner that supports multiple search strategies and finds multiple acyclic plans (e.g., Top-k or Top-x%). In this experiment, we weighted each of the three objectives (min distance, max concealment, and max order consistency) equally, and extracted the Top-30 plans. However, the Top-30 plans contained very minimal diversity, and were closely related to the optimal minimal distance route. Using equal weighting, the next meaningfully different route was discovered at solution #987 (17 moves), followed by #994.

We can visualize the trade-offs between the various objectives by comparing the scores of the individual objectives along with the combined weighted score. These are plotted as Figure 5 after being normalized for comparison and inverted (better scoring is shown as larger, although the planner is actually minimizing costs). It shows three unique options and a tradeoff. In top cases, prioritizing distance resulted in lower cover and order scores. Prioritizing cover and order resulted in lower distance and weighted scores. There were also some routes feasible with high distance and cover at the cost of order and weighted scores.

In order to quantify the trade-offs, a correlation matrix can be run on the scores for the solutions (Table *4*). There is a strong positive correlation with the objective to minimize distance and the equally weighted score, showing that this objective tended to override the others. Since time in any other feature depended on total time and because all the routes score near the optimal route distance, it is not surprising that solutions that scored non-optimally for this objective didn't appear until after many solutions were generated.

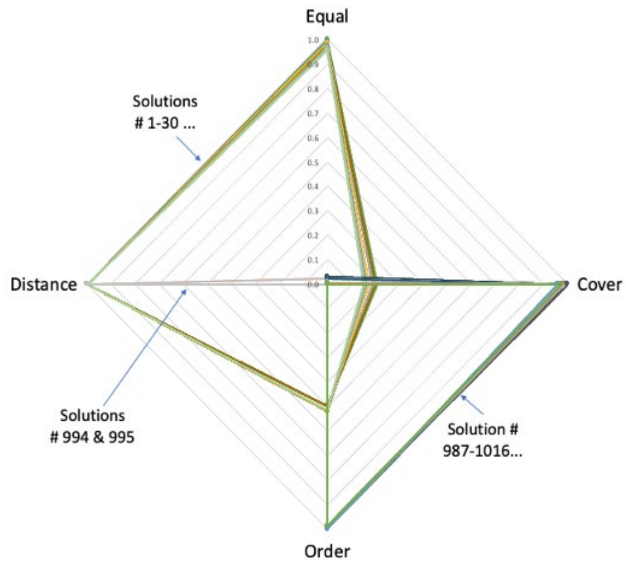


Figure 5. Equally Weighted Search - Radar chart showing scores (larger is better) of Top-30 solutions (1-30) as well as solutions 987-1016. These later solutions show the first meaningfully diverse trade-off of objectives and they are sparse.

	Equal	*Cover*	*Order*	*Distance*
Equal	1.000			
Cover	-0.906	1.000		
Order	-0.767	0.964	1.000	
Distance	0.936	-0.996	-0.943	1.000

Table *4*. Correlation matrix for objectives in the route planning problem when optimized with equal weights.

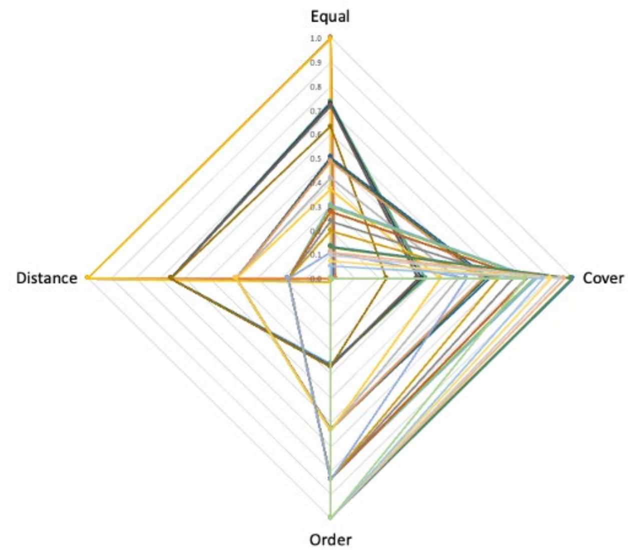


Figure 6. Pareto Search - Radar chart showing scores (larger is better) of 30 solutions on the Pareto Front. This shows that no solution dominates all the others and exposes a wider range of meaningfully diverse solutions with trade-offs of objectives.

	Equal	*Cover*	*Order*	*Distance*
Equal	1.000			
Cover	-0.928	1.000		
Order	-0.988	0.973	1.000	
Distance	0.988	-0.974	-1.000	1.000

Table 5. Correlation matrix for solutions on Pareto Front.

Routes on the Pareto Front of Objectives

To generate more diverse solutions, we developed a search strategy focused on discovering solutions on the pareto front for all of the objectives. Similar to (Choudhury 2016), we developed a strategy for directing state-space search across all of the objective dimensions of the space and testing for non-dominated solutions. We generated 30 solutions on the Pareto Front, but unlike the weighted case, we cannot call these “top” solutions. For comparison, we directed the planner to compute metrics for an equally weighted objective without including it in the optimization. As shown in Figure 6, the resulting solutions show a significantly more diverse (in terms of metric space) set of solutions, as intended.

The correlations for the solutions on the Pareto front are consistent with, but slightly stronger than, the previous method (Table 5). One significant benefit was that the variation in solutions was increased, and all 30 solutions produced were useful compared to the first nearly 1,000 solutions showing little variation. Moreover, many solutions had middling scores on the equally weighted objectives but represented routes that had a reasonable trade-off across objectives for the given situation.

Conclusions and Future Work

In this paper, we compared *a priori* and *a posteriori* strategies for directing the search for both position and route elements of military course of action planning. We used methods that are both practical and doctrinally explainable in After Action Review. For each step, we developed and evaluated a diverse set of feasible solutions representative of the potential trade-offs between multiple, oft competing, objectives, and quantified them as correlations in the metric space. In directing optimization towards discovering solutions on and near the Pareto Front, we identify situationally relevant trade-offs a decision-maker might consider.

The next steps for this research are to leverage the range of trade-offs in both creating automated behaviors (e.g., for a simulated force), but also in better recognizing and anticipating the range of potential behaviors in others (e.g., a human trainee). In this latter case, incorrectly assuming a set of *a priori* weights across multiple competing objectives could result in misinterpretation of observations (e.g., bias). Conversely, learning weights for competing objectives repsentated in trainee behaviors through observations may prove valuable in better understanding behaviors of trainees.

Acknowledgements

This material is based upon work supported by the US Army DEVCOM SC STTC under contract No. W912CG-21-C-0018. Any opinions, findings and conclusions, or recommendations expressed are those of the author(s) and do not necessarily reflect the views of the DEVCOM SC STTC.

References

Argenta, C., Kewley, R., and Brawnder, K. Behaving Like Soldiers: A Multi-Agent System Approach to Course of Action Planning for Simulated Military Units, *Proceedings of Florida Artificial Intelligence Research Society (FLAIRS-35, 2022)*.

Choudhury, Shushman & Dellin, Christopher & Srinivasa, Siddhartha. (2016). Pareto-optimal search over configuration space beliefs for anytime motion planning. 3742-3749. 10.1109/IROS.2016.7759551.

Hart, P., Nilsson, N., & Raphael, B. (1968). A Formal Basis for the Heuristic Determination of Minimum Cost Paths. IEEE Transactions on Systems Science and Cybernetics, 4(2), 100–107. https://doi.org/10.1109/tssc.1968.300136

Headquarters Department of the Army 2019, *Army Doctrine Publication 6-0 – Mission Command: Command and Control of Army Forces*. Washington, D.C.: The Army Publishing Directorate.

Headquarters, Department of the Army (HQDA), 2022. *Field Manual 3-0: Operations*. Washington, D.C.: The Army Publishing Directorate.

McGroarty, C and Gallant, S 2022. Expectation Management of AI for Automated Military Behavior Representations, *Proceedings of Florida Artificial Intelligence Research Society (FLAIRS-35, 2022)*.

Narayanan, P. et al, *First-Year Report of ARL Director's Strategic Initiative (FY20–23): Artificial Intelligence (AI) for Command and Control (C2) of Multi-Domain Operations (MDO),* ARL-TR-9192, US Army Development Command – Army Research Laboratory.

Nebro, A, Durillo, J., Luna, F., and Alba, E., 2009. Evaluating New Advanced Multiobjective Metaheuristics, Optimization Techniques for Solving Complex Problems, Alba et. al. eds., ch. 5.

Ngatchou, P., Zarei, A., and El-Sharkawi, M., 2005. Pareto Multi Objective Optimization, *Proceedings of the 13th International Conference on, Intelligent Systems Application to Power Systems.*

Owens, K, Gupton, K, Hellman, K, and Goldberg, B 2022. Re-Thinking the Tactical Small Unit Synthetic Training Model, *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC 2022)*.

Roper, D.S. 2018. Regaining Tactical Overmatch: The Close Combat Lethality Task Force. Institute of Land Warfare. Retrieved Feb 8, 2023 from https://www.ausa.org/sites/default/files/publications/SL-18-2-Regaining-Tactical-Overmatch-The-Close-Combat-Lethality-Task-Force-0.pdf.

Russell, S. J., & Norvig, P. (1995). Artificial intelligence: a modern approach. Englewood Cliffs, N.J., Prentice Hall.

Scales, R.H. 2019. Tactical Art in Future Wars. Retrieved Feb 8, 2023 from https://warontherocks.com/2019/03/tactical-art-in-future-wars/.