Co-evolutionary Information Gathering for a Cooperative Unmanned Aerial Vehicle Team

Jean Berger
DRDC Valcartier
Quebec City, Canada
jean.berger@drdc-rddc.gc.ca

Jens Happe, Christian Gagné
MacDonald, Dettwiler & Associates Ltd.
Richmond, B.C., Canada
jhappe@mdacorporation.com

Martin Lau
Dept of Computer Science
University of British Columbia
Vancouver, Canada

Abstract - *Persistent surveillance and reconnaissance tasks in mobile cooperative sensor networks are key to constructing recognized domain pictures over a variety of civilian and military problem instances. However, efficient information gathering for a task such as target search by a team of autonomous unmanned aerial vehicles (UAVs) still remains a major challenge to achieve system-wide performance objective. Given problem complexity, most proposed distributed target search solutions so far consider simplifying assumptions such as predetermined path planning coordination strategy with implicit communication and ad hoc heuristics, and severely constrained resources. In this paper, we extend previous work reported on multi-UAV target search by learning resource-bounded multi-agent coordination, involving explicit action control coordination. The approach first relies on a new information-theoretic co-evolutionary algorithm to solve cooperative search path planning over receding horizons, providing agents with mutually adaptive and self-organizing behavior. The anytime algorithm is coupled to an extended information-sharing policy to periodically exchange world-state information and projected agent intents. Preliminary results show the value of the proposed approach in comparison to existing techniques or methods.*

Keywords: Learning, multi-agent, coordination, co-evolution, heuristics, unmanned aerial vehicles.

1 Introduction

Efficient construction of a recognized air picture (RAP) for military local area surveillance and reconnaissance missions is often critical to ensure and maintain situational awareness. In many cases, given the low cost and risk associated with resource utilization, the related RAP process increasingly relies on a team of mobile sensor agents or unmanned aerial vehicles (UAVs) (both terms are used interchangeably) to perform cooperative search, closing the gap between information need and information gathering. Early work on related search problems emerges from search theory [1], [2]. Proposing a mathematical framework and models leading to analytical solutions for simple static formulations, most efforts have progressively been devoted to algorithmic contributions to handle more complex dynamic problem settings [3], [4]. However, reported work for these problems mainly focused on centralized search while assuming search effort to be infinitely divisible between cells, making it difficult to solve realistic path planning problems [4] in cases where cell target containment probability is sparsely distributed. Search theory solutions mostly relate to the effort allocation decision problem rather than to path construction. Robot motion planning alternatively explored search path planning, primarily providing constrained shortest path type solutions for coverage problem instances [5]-[7]. Even though cooperation of multiple UAVs already proved its value over individual vehicles operating independently [8], [9], this paper further focuses on constrained cooperative search path planning for multiple UAVs in a dynamic uncertain environment. In this setting, teammates must self-organize, autonomously manage their own resources, and coordinate their behavior to achieve a commonly shared system-wide global objective. Typical decision problem formulations and solutions are presented in [10]–[15]. Recent extensions to this work further address the critical information-sharing dimension of the cooperative search planning problem [16]–[19]. These problems capture to various degrees some simplified decentralized partially observable Markov decision process combining communication (informed-sharing) and control decisions (COM-DEC-POMDP) [20], [21] which has proven to be NEXP-complete [22]. Facing the curse of dimensionality, exact problem-solving methods to sequential decision problem formulations, such as dynamic programming, are generally not practical, paving the way to the development of efficient heuristic and approximate methods. Explicit solutions proposed for multiple-UAV cooperative search path planning are numerous. Some early approaches simply reduce computational complexity by relaxing some hard constraints to keep the problem manageable. Presuming continuous full state observability by the team while confining search to particular regions of the solution space [23] represents such an example. Other procedures often assume unlimited computational and communication resources. As a result, sub-optimal or possibly unsuitable solutions are generated. Methods inspired from search theory propose procedures based on branch and bound or A* types of techniques, but the determination of good heuristics to compute tight bounds for long-term solution quality estimation largely remains difficult [4]. Liao et al. [16] recently proposed a search path planning approach combining a cooperative path planning control solution coupled to a specific predetermined information-sharing policy. The proposed information-sharing policy considers a unicast communication scheme and small-world assumption. However, the proposed constrained solution and alternative approaches reported in [23]–[26] do ignore path planning coordination that, though constrained, could take place through explicit communication of intents (agent path plans). The approach also assumes unbounded agent communication bandwidth over a local neighborhood to mutually share historic observations or beliefs with close neighbors. In that respect, it simply assumes sufficient time to exchange the possibly large number of observations about world state beliefs, but not enough time to engage in a path planning coordination process implying explicit communication (e.g. negotiation). Current solutions proposed so far for UAV teams paid little attention to path planning coordination involving explicit communication under stringent spatial and temporal constraints, and ultimately focused on the construction of single-agent solutions, limiting system robustness.

This paper presents a co-evolutionary information gathering approach in cooperative unmanned aerial vehicle teams to carry out target search. It is primarily inspired from the information-theoretic framework reported in [23]–[26] and the recent contribution presented in [16] to address information sharing. It extends previous work reported on multi-UAV target search by learning resource-bounded multi-agent coordination for a new problem, considering an open-loop with feedback decision model with a rolling horizon, multiple objectives, heterogeneous agents, limited computational resources and communication bandwidth, as well as communication cost, in a time-constrained uncertain environment. It concurrently deals with multiple constraints, departs from predetermined control search plan policy based on implicit or passive plan coordination [26], and proposes a framework to construct joint path plans providing team flexibility and adaptation by co-evolving multiple agent behaviors simultaneously while mitigating communication needs and cost. The main contribution lies on a new information-theoretic co-evolutionary algorithm to solve cooperative search path planning over receding horizons, providing agents with mutually adaptive and self-organizing behavior. The anytime algorithm is coupled to an extended information-sharing policy to exchange world-state information and projected agent intents.

The remainder of the paper is structured as follows. Section 2 first presents the problem definition, describing the main characteristics of a new cooperative information-gathering problem involving multiple UAVs to carry out target search. Then the main solution concept for the targeted problem is introduced in Section 3. A two-step decomposition scheme to sequentially achieve state estimation through information-sharing and cooperative path planning is outlined. Section 4 reports and discusses preliminary computational results comparing the value of the proposed method to alternative techniques. Finally, a conclusion is given in Section 5.

2 Problem

2.1 Description

We consider a hierarchical multi-objective problem in which a team of heterogeneous UAVs cooperatively searches stationary targets in a bounded environment over a given time horizon. The first objective is to maximize information gain or equivalently to minimize uncertainty or entropy about target occupancy over the grid, the second consists in minimizing target discovery time, and the third aims at minimizing resource utilization, namely, energy consumption. The proposed hierarchical objective structure refers to lexicographic ordering, ranking solution quality along the described objectives, respectively in that order. Modeled as a grid, the search environment defines a two-dimensional cellular area composed of N cells, populated by non-cooperative stationary targets and threats, which are assumed to occupy a single cell each. Contrarily to threats, the number of targets and their respective locations are initially unknown. Based on prior domain knowledge, individual cells are characterized by an initial probability of target occupancy, which defines an agent's cognitive map. The target occupancy probability is assumed to be independent between cells. Given an initial team configuration, autonomous UAVs synchronously explore the environment, acting as stand-off imperfect sensors gathering observations while periodically exchanging state and plan information with one another through peer-to-peer (unicast) communication. UAVs perfectly and synchronously share information (observations, intents) with neighbor agents during each discretized episode or time step (visit). Information-sharing is subject to limited on-board computational resources and range-constrained communication. Vehicles are assumed to fly at a constant velocity and at slightly different altitudes to avoid colliding with each other. Cooperative search consists in jointly constructing agent path plans to minimize team uncertainty (entropy) over cell target occupancy. As the system is distributed, each agent (vehicle) must continuously build and update its own cognitive map by making sensor observations and by exchanging information with teammates while aligning its behavior toward reaching global team objective. A UAV's cognitive map refers to a knowledge base capturing local environment state representation, reflecting target occupancy belief distribution, positions and orientations of neighboring agents, its own direction and position, resource level, sensor observations and their sources (observing agent), as well as a past communication log with other agents. It is also assumed that each agent has prior knowledge about its teammates (e.g. sensor observation models, maximum communication range and other properties). A typical agent cognitive map at a given point in time is illustrated in Figure 1.

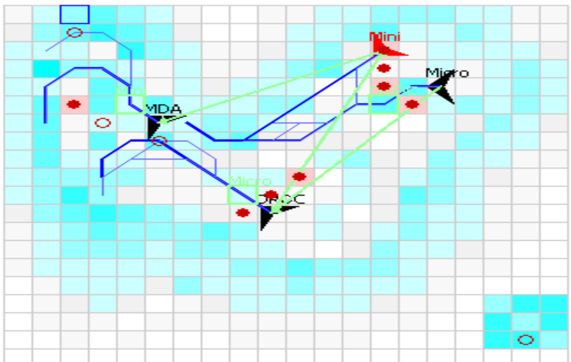


Figure 1. Agent uncertainty/cognitive map at time step *t*. Local agent beliefs are displayed through multi-level shaded cell areas. Projected agent plans are represented as possible paths and peer-to-peer communications as straight lines. Filled circles refer to discovered targets.

2.2 Observation model

The observation model governing a UAV sensor's perception accounts for partial world state observability. During each time step, a UAV visits a cell searching for target occupancy. A sensor reading z_t at time *t* may be positive or negative and is governed by an observation model, which accounts for uncertainty through conditional probability of detection and false alarm given cell target occupancy/vacancy state (*X*=1/0):

z_t : observation of cell occupancy at the end of period *t*: {positive=1, negative=0}

$p_c = p(z_t = 1 \mid X = 1)$ probability of correct observation

$p_f = p(z_t = 1 \mid X = 0)$ probability of false alarm

In addition to probabilistic outcomes introduced by imperfect sensor readings, vehicles have limited sensing ranges within which they can perceive and recognize neighboring agents.

2.3 Bayesian filtering

Based on the latest observation, local cell target occupancy beliefs (*p(X=1)*) can be modified using Bayesian filtering for data fusion:

$$p_t(X \mid z_t) = \frac{p(z_t \mid X) p_{t-1}(X)}{p(z_t)} \qquad (1)$$

where

$$p(z_t) = \sum_{x \in \{0,1\}} p(z_t \mid X = x) p_{t-1}(X = x) \qquad (2)$$

p_{t-1} and p_t refer to prior and posterior cell target occupancy probability (belief) respectively. Sensor readings may be shared with neighboring agents after each observation episode (cell visit) to further update local beliefs and progressively build a more consistent cognitive map.

2.4 Path planning

A vehicle in a configuration state dictated by a given position (cell location), and a specific orientation (*N,S,E,W,NE,SE,SW,NW*) makes a decision at each episode about its next visit. Decisions are limited to three possible moving directions with respect to its current heading, namely, *ahead, right, left*. The primary goal consists in planning base-level control action moves to minimize entropy (target occupancy uncertainty) over the entire grid. The entropy function *E* is borrowed from information theory [27]:

$$E = -\sum_{x \in \{0,1\}} p(x) \log_2 p(x) \qquad (3)$$

where *p* refers to the current probability/belief of cell target occupancy. A cell entropy of 0 (1) means absolute certainty (total uncertainty) about occupancy or vacancy. UAVs are subject to resource-bounded reasoning due to limited computational resources, constrained communication and temporal constraints imposed by episode duration.

2.5 Communication

The agent team primarily behaves as a particular mobile ad hoc network. Vehicles have self-localization capability, can recognize neighboring agents, and rely on a unicast or peer-to-peer communication scheme based on perfectly reliable communication channels. Accordingly, the model assumes range-limited communication restricting neighborhood interaction, but sufficient bandwidth to support neighbor exchanges during a single time step. It should be noticed that agent heterogeneity makes the neighborhood relationship asymmetric. Agent communications with neighbors take place concurrently delivering/receiving messages on separate channels in parallel. Encoded as messages, communication decisions translate into observation streams, beliefs and/or intents to be shared. Based on the aforementioned small-world assumption, we also assume instantaneous message-passing (no network latency), ignoring the impact of routing considerations. Energy cost supporting information exchange is quadratic in terms of the distance *r* connecting two agents ($\alpha r^2 + \beta$).

3 Solution concept

The proposed solution to cooperative multi-UAV search path planning relies on a co-evolutionary information gathering approach. It extends previous work reported on multi-UAV target search [23]-[25], [16], [26], by learning resource-bounded multi-agent coordination considering an open-loop with feedback decision model over a receding horizon, multiple objectives, heterogeneous agents, limited computational and communication resources, as well as communication cost, in a time-constrained uncertain environment. It proposes a framework to construct joint path plans providing team flexibility and adaptation by co-evolving multiple agent behaviors simultaneously while mitigating communication need and cost. The new information-theoretic co-evolutionary algorithm solves cooperative search path planning over receding horizons, providing agents with mutually adaptive and self-organizing behavior.

Each time step, an agent's decisions rely on a sequential two-stage decomposition process, namely, information–sharing (past observations) and planning. The time interval Δt_t is divided in two time subintervals, accounting for information-sharing and planning respectively, as shown in Figure 2.

Latest observations are first exchanged between an agent and its close neighbors, the remaining time being devoted to cooperative planning, which in turn is subdivided into cycles, involving multiple planning iterations coupled to asynchronous periodic communication of intents. In order to reinforce agent synchronization, we assume that each process stage has a

Information-sharing communication – state/observationS	Planning planning iterations & periodic intent sharing
← $^{s}\Delta t_{c,t}$ →	← $\Delta t_{pl,t}$ →

←Action execution: move to the next cell and observe→

Time step t (episode) with duration Δt_t

Figure 2. A two-stage decision process

constant predetermined duration, which imposes temporal constraints on information state communication and planning. Decision-making for the current episode is occurring concurrently with the execution of the action planned in the previous episode. Therefore, during time interval Δt_t, the vehicle executes a previously planned action, moves to the next cell, and makes an observation. The process can be summarized as follows:

t=1
For agent i= 1..n
  Initialize Population(i)
Repeat -- agent search path planning behavior
  Control action execution (planned at t-1)
  Information_Sharing ($t(^{s}\Delta t_{c,t})$) – Stage I
  Path_Planning ($t(\Delta t_{pl,t})$) – Stage II
  Observation and cognitive map update
  End of episode t; t=t+1
Until (end of search mission horizon: t=L)

Section 3.1 gives further details on the internal processes associated with path planning as well as population initialization.

3.1 Cooperative search path planning

An open-loop agent solution to a multi-objective problem subject to limited computational resources and communication constraints is gradually constructed at each episode and progressively extended over a receding T-step horizon, while adjusting its path plan based on additional feedback. Episodic search path planning relies on co-evolution to learn agent coordination. Agents evolve their own population of individuals while sharing information about neighbor agent intents. Individuals are represented as chromosomes encoding for a given time step, a feasible path plan expressed as a sequence of intended control actions (physical moves a_{t+1} ... a_{t+T}) to be executed over a specific time horizon T (Figure 3).

a_{t+1}	a_{t+2}	a_{t+3}	…	a_{t+T-1}	a_{t+T}

Figure 3. An individual path plan representation (genotype) at time step t.

Agents co-evolve their own path plan individuals through natural selection, recombination and mutation mechanisms over successive generations while periodically (T_{GA} or *max_gen* generations) exchanging their best individual with neighbor agent populations. An individual fitness is determined by combining(coalescing) its own local path plan along with the latest best-known neighbor agents' plans and evaluating the resulting joint plan. Here the fitness evaluation of a plan is based on the agent's interaction with other agents. The cycle is then repeated until the end of the planning period. As a result, the first action of the best computed path plan is executed at the next time step. The agent co-evolutionary anytime procedure can be summarized as follows:

Path_Planning(t)
  Adjust/update population individuals to reflect the latest agent's decision (current move)
  Repeat
    gen=0
    Repeat for each new generation
      For all new "best plan" message received
        adjust individual's fitness value
      Evolve Population - build a new generation
        generate λ new offspring using genetic operators (selection, recombination, mutation)
        evaluate fitness of new individuals
        eliminate the λ worst individuals of the expanded population
      gen=*gen*+1
    Until (end of a cycle (T_{GA}) or *gen* = *max_gen*)
    intent-sharing (send current 'best plan' to neighbors)
  Until (end planning period)
  Return (best computed path plan from Population)

The first step consists in adjusting the population members emerging from the last episode to account for current move execution, conserving feasible (consistent) individual path plans only, and generating randomly new individuals (path plans with T control actions) to maintain population size. Newcomers could be partly created from best known heuristic methods as well, guaranteeing minimal solution quality. Feasible path plans are simply shifted by one step to reflect the current time, and a new action for time step T (plan horizon) is drawn randomly from the three possible actions, with a probability proportional to its related information gain. The outer loop captures multi-agent coordination learning and intent-sharing between neighbor agents. The underlying evolutionary approach and its key components are mainly outlined in the inner loop. The steady-state evolutionary algorithm consists for each generation in expanding the population by λ offspring using genetic operations, and then removing the worst λ individuals to restore normal size. Recombination and mutation operators are sequentially applied with probability p_{Xover} and p_{Mut} ($p_{Xover} + p_{Mut} = 1$), respectively, until all λ new individuals are generated. Probability parameters are generally determined to balance search intensification and diversification. The process is repeated until a convergence criterion/condition is met (e.g. maximum number of generations or a threshold in solution quality improvement). Related fitness definition and genetic operators are further described below.

The initial population of path plan individuals is generated by selecting control actions randomly. For each future time step over the horizon T, probabilistic action selection is proportional to the information gain over all one-step possible moves. Best known heuristic methods can be partly exploited to generate good individuals.

3.1.1 Fitness:

Fitness reflects an individual's propensity to reproduce. The individual's fitness should be determined and ranked according to the hierarchical objective structure proposed in Section 2.1 and based on lexicographic ordering: maximize information gain, minimize target discovery time and energy consumption. As dissimilar fitness values over population individuals are likely, we approximate the global fitness function along entropy minimization (or information gain maximization) to more easily rank individual scores. This approximation makes sense to the extent that target discovery rate is somewhat correlated to entropy and that resource utilization is mainly controlled by the predetermined information-sharing policy.

Fitness is defined in terms of a weighted combination of expected information gains (differential entropy contributions) capturing shared individual path plan and neighboring agents current "best plan" (neighbors' population best-known individual) respective expected rewards, while ignoring possible additional benefits that could have resulted from intermediate explicit outcome communication or exploitation. An agent's individual reward refers to the local information gain or entropy reduction expected by executing its path plan over the next plan horizon T. Subject to range and temporal constraints, communication on respective neighboring agents best plan (intents) take place periodically among agents after a certain number of generations, to update agent individual's fitness value. That period can be initially adjusted to get a lower bound on the number of messages to be exchanged with some neighbors during the planning phase. It should be mentioned though, that mutual information-sharing occurs asynchronously given heterogeneous communication constraints and continuous plan computation, meaning that individual fitness values are progressively updated during the next planning algorithm execution cycle between each new population generation, on the arrival of new "best plan" messages in the agent buffer. The nominal fitness function for an individual *ind* in agent population i characterized by a local path plan or sequence of actions $\{a_{it}\}$ (resulting in positions/expected information gain value pairs $\{y_{it}(a_{it}), {}^i g(y_{it})\}$) along with respective neighbors' best-known path plan and information gain ($\{y_{jt}, {}^j g(y_{jt})\}_{j \in Neigh(i)}$) over the next T time steps, is defined as follows :

$$fitness_{ind \in Pop(i)}(\{y_{jt}, {}^j g(y_{jt})\}_{j \in Neigh(i) \cup \{i\}}) = \sum_{j \in Neigh(i) \cup \{i\}} {}^j R$$
$$= \sum_{c \in G} \sum_{j \in Neigh(i) \cup \{i\}} \frac{{}^j g(c)}{\sum_{k \in Neigh(i) \cup \{i\}} {}^k g(c)} {}^j g(c) \quad (4)$$

$${}^j g(c) = {}^j E_0(c) - {}^j \overline{E}_T(c) = {}^j \Delta E(c) \quad (5)$$

$${}^j \overline{E}_T(c) = \sum_{\{z_{jt}(c)\} \in \{0,1\}^T} \left[\prod_{t=1}^{T} p(z_{jt}(c) | y_{jt}(a_{jt}), o_{jt}(a_{jt})) \quad {}^j \varepsilon_T \right] \quad (6)$$

$${}^j \varepsilon_T = {}^j E_T(p_T(X(c) = 1 | \{z_{jt}(c)\}, \{y_{jt}, o_{jt}\})) \quad (7)$$

c: cell element of the grid: $c \in G = \{1,2,...,N\}$

T: time horizon

i: agent team member $i \in \{1,2,...,n\}$

a_{it} : action of agent i executed during time interval t
$a_{it} \in \{ahead, right, left\}$

$\{a_{it}\}_{1..t}$: path plan of agent i over history $1..t$

y_{it} : position (cell element number) of agent i over time interval t as a result of action a_{it}

o_{it} : orientation of agent i over time interval t as a result of action a_{it}

$z_{it}(c)$: observation outcome of cell c by agent i at the end of time interval t, $z_{it}(c) \in \{0,1\}$

$\{z_{it}\}$: sequence of observations by agent i over history $1..T$

$p(z_{it}(c))$: probability to observe outcome $z_{it}(c)$

$p_t(X(c) = 1)$: belief of cell occupancy of cell c at the end of time interval t

${}^{i}E_T(p_T(X(c)=1|\{z_{it}(c)\},\{y_{it},o_{it}\}))$: agent *i* entropy of cell *c* at *t=T* given the sequence of actions $\{a_{it}\}$ and observations $\{z_{it}\}$ over history 1..*T*

Given that the global team entropy is unknown, it can be noticed from the fitness function that interfering agents (planning to visit the same cell) share information gain proportionally to their relative contribution to better account for the reduced team information gain. For example, two agents exploring the same cell with identical anticipated benefits will ultimately see their respective cell reward reduced by half. The agent's information gain reflects the local entropy reduction expected by executing the open-loop agent's path plans over all possible *T*-step scenarios. Local information gain per cell for agent *j* (${}^{j}g(c)$) can be expressed as the difference between current entropy (${}^{j}E_0(c)$) and expected entropy (${}^{j}\overline{E}_T(c)$) resulting from plan execution. For a given plan, ${}^{j}\overline{E}_T(c)$ can be computed more efficiently based on the specific planned cells to be visited (O(*nT*)) rather than explicitly enumerating all possible histories or sequence of observation/visit outcomes (2^{nT}).

3.1.2 Selection:

In order to balance and control selection pressure for breeding purposes, fitness values are sorted out and normalized using a linear ranking scheme [28], scaling respective values based on rank. This turns out to be advantageous in avoiding premature solution convergence or random exploration. The technique is particularly useful for a population of individuals implicitly presenting small relative (nominal) fitness variability (or alternatively, clustered fitness distribution), emphasizing the selection of the fittest individual without neglecting other population members. Accordingly, an individual *ind* from population *Pop* (size) with a fitness value ranking $rank^{th}$ is scaled to a new value as follows:

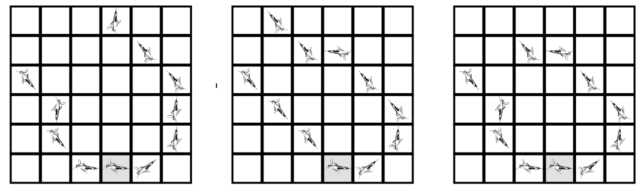
$$fitness'_{ind} = \max - \left[(\max - \min)(rank-1)/(Pop-1)\right] \quad (8)$$

with max=1.6 and min=0.4 .

Individual selection for mating purposes is ensured following a fitness-proportional scheme [29]. The larger the fitness value, the larger the probability to be selected. Using the above fitness linear ranking, the selection scheme allows to more frequently discriminate the best individual over the worst, by a ratio of max/min.

3.1.3 Recombination:

This genetic operation recombines chromosomes from two selected parents in order to create a child. The proposed recombination operator X_path breeds two parent individuals sharing a compatible crossover point and generates an offspring by connecting together head and tail path segments inherited from both parents respectively, truncating control actions when the chromosome length exceeds the planning horizon *T* (see Figure 4) or appending missing control actions using a greedy method (selecting moves with maximum one-step information gain) to complete the solution when necessary. A second child can be generated in the same way by swapping parents. The operator mainly relies on the key notion of parent compatibility. In its simplest form, two parents are compatible if their paths cross each other at least once, while exhibiting dissimilar sub-routes prior and posterior to the crossing point (avoiding parent duplication). A potential mating point can be computed in $O(T^2)$ time. Should many and/or redundant mating points exist, the earliest crossover point from the first parent would be selected for the recombination operation.

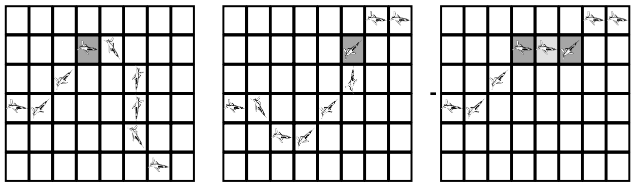


Parent 1 Solution **Parent 2 Solution** **Child Solution**

Figure 4. Crossover operator X_path mating Parent 1 and 2 to generate a new child solution. Parent trajectories are shown to intersect at a cross-over point depicted by the shaded cell. The last control action inherited from Parent 2 is deleted to maintain solution consistency.

However, unless agent orientations for the crossing point are similar for both parents as shown in Figure 4, which is a particular case, local repair of the offspring path solution is generally necessary to make it feasible. In effect, state inconsistency occurs as the child inherits agent orientation from the first parent and remaining control actions from the second, which may violate kinematics constraints and render the move at the crossover point impossible or illegal. This problem can be circumvented by generalizing the definition of individual compatibility introduced earlier in its most restricted form, as depicted in Figure 4. Generalizing path crossing requirements, compatibility is stated in terms of the number of moves required to locally modify a given solution path toward concurring and ultimately converging to an alternate solution trajectory. This number of moves is a distance measure between two solutions. A solution is *d*-compatible with another if it can concur toward the other in at most *d* moves: The smaller the distance, the larger the compatibility. Figure 5 illustrates a typical recombination operation involving 2-compatible individuals. Accordingly, selected candidate individuals are examined for 2-compatibility condition until a match arises, at which point the recombination operator is activated.

The approach tends to make the child inherit as much as possible the original path structure of both parents, minimizing path segment distortion due to the crossover process. As a special case, Figure 4 pictures an example where Parent 1 is 0-compatible with Parent 2, mainly pre-



Parent 1 Solution **Parent 2 Solution** **Child Solution**

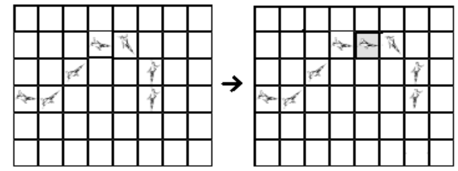
Figure 5. Crossover operator X_path recombining Parent 1 and 2. Parent 1 is 2-compatible with Parent 2: it requires the insertion of a single move connecting inherited parent path segments to generate a feasible child. The inserted move is pictured as the additional shaded cell in the emerging child solution.

serving the structure of path segments inherited by the child solution. Therefore, by bounding the distance (d_{max}) separating two selected candidate individuals when exploring suitable recombination, we reduce the cost otherwise associated with unfeasible child solution generation and local repair. Candidate individuals are sequentially visited for d-compatibility (d=0,...d_{max}).

Dissimilar candidate individuals d-compatibility computation is in $O(T^2 \log(d) + d^4 T)$, which is acceptable if d is suitably small (e.g. d<5). Computational complexity can be further reduced by pruning the search space, such as exploring partial parent path segments only, or parents exhibiting path centroid (average path position) proximity. Forward move projection span (reachability space) of possible trajectories over a horizon d is first generated for each move (position-orientation pair) of Parent 1 with index t (t<T–d). Reachable cells and possible orientations from the original move can then be sorted out to retrieve and match any Parent 2 current moves. Projected move trajectory construction and specific targeted Parent 2 move exploration can be processed in any order. The dual notion of "backward" move projection from Parent 2 could be exploited as well to match a specific targeted Parent 1 move, achieving similar results. A hybrid approach might even match elements from the forward projection span from one parent move to an alternate backward move projection element from the other parent to reconstruct a child solution.

3.1.4 Mutation:

Mutation is a natural evolution process modifying some individual's genes more or less frequently. Two mutation operators are proposed, namely, M_path and M_path_local_repair. M_path first consists in selecting a specific move (index t>2) composing a path solution, modifying it randomly with an alternate action and then reconstructing a feasible remaining solution from that point, as shown in Figure 6. From the altered move (gene at index t) on, the last portion of the path (chromosome) is generated, by consecutively and randomly adding new moves (genes) until a complete solution emerges. Move probability selection is proportional or biased toward the best expected one-step information gain.



Solution **Mutated Solution**

Figure 6. Mutation: a move from an individual solution is randomly selected and mutated. A new solution is then reconstructed from that point (shaded move).

The M_path_local_repair mutation operator randomly removes a short path segment (few consecutive moves, at least two steps from the path endpoint) from a solution and builds an alternate path segment fragment to locally repair or bridge the disconnected components. d-compatibility between disconnected segments (disconnected component end points) is exploited to reconstruct a slightly different path solution, by minimizing changes required to preserve as much as possible the structure of the parent.

4 Preliminary results

A preliminary computational experiment has been conducted to illustrate the proposed framework. From a limited simulation sample for a given limited scenario involving a team of three UAVs, it shows the value of explicit team coordination comparing the co-evolutionary approach to an alternative method derived from a well-known heuristic [16] involving self-interested agent behavior in which there is no coordination. In the self-interested scheme, agents are planning their search path independently using a greedy method, which consists in selecting the next move based on maximum information gain expected over a one-step time horizon. A typical result for a 3-UAV team with limited communication range is shown in Figure 7 presenting post-processed

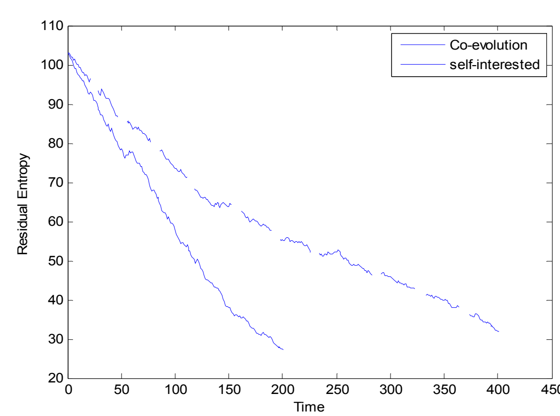


Figure 7. Differential entropy displaying the value of coordination. The lower the entropy, the better the solution.

team entropy over time. The differential entropy in the graph clearly shows the advantage of co-evolutionary team behavior over a self-interested greedy attitude.

5 Conclusion

A new cooperative search path planning problem for multiple heterogeneous UAVs subject to computational and communication resource constraints, has been addressed. Towards solving this problem and closing the gap between information need and information gathering, an information-theoretic co-evolutionary framework has been proposed. The approach extends previous work reported on multi-UAV target search by learning resource-bounded multi-agent coordination, considering an open-loop with feedback decision model and multiple objectives. It solves cooperative search path planning over rolling horizons, computes multiple solutions/options, and exhibits adaptive and self-organizing team behavior. The anytime algorithm is combined with an information-sharing policy to communicate world-state information and intents on a periodic basis. Early results show the value of explicit team coordination.

Future work will investigate strengths and weaknesses of the proposed approach through a comparative performance analysis for a variety of heuristic methods and conditions. Then, the approach will be further extended to explicitly address cooperative information-sharing under limited communication bandwidth.

References

[1] Benkoski, S., Monticino, M., and Weisinger, J., 1991. A survey of the search Theory Literature, *Naval Research Logistics*, 38.

[2] Stone, L. 1989. What's happened in search theory since the 1975 Lanchester Prize? *Operations Research*, **37** (3).

[3] Hohzaki, R., Iida, K. 1995. Optimal search plan for a moving target when a search path is given. *Math. Japonica*, 41(1).

[4] Lau H. 2007. Optimal Search in Structured Environments, *PhD Thesis*, University of Technology, Sydney 2007.

[5] Rekleitis, I. et al., 2008. Efficient Boustrophedon Multi-Robot Coverage: an algorithmic approach, *Ann Math Artif. Intell.* 52:109–142.

[6] Agmon, N. et al., 2008. The giving tree: constructing trees for efficient offline and online multi-robot coverage, *Ann Math Artif. Intell.* 52:109–142.

[7] Wagner, W. et al., 1999. M. Distributed covering by ant-robots using evaporating traces, *IEEE Trans. Robot. Autom.*, vol. 15, no. 5.

[8] Chandler P., Rasmussen, S., Pachter M., 2000. UAV Cooperative Path Planning , *Proc AIAA Guidance, Navigation, and Control*.

[9] Chandler P., Pachter M., 2001. Hierarchical control for autonomous teams, *Proc AIAA Guidance, Navigation, and Control*.

[10] Mathews, G., Durrant-Whyte H. 2006. Scalable Decentralised Control for Multi-Platform Reconnaissance and Information Gathering Tasks. In *9th International Conference on Information Fusion.*

[11] Jin, Y., Liao, Y., Minai, A., Polycarpou, M. 2006. Balancing search and target response in cooperative unmanned aerial vehicle (UAV) Teams *IEEE Trans on Sys Man and Cybern. Part B*, 36(3).

[12] Flint, F., E. Fernandez-Gaucherand, E., and M. Polycarpou, M., 2004. Efficient Bayesian Methods for Updating and Storing Uncertain Search Information for UAV's. *Proc 43rd IEEE Conf on Decision and Control*.

[13] Beard, R.,. McLain, T., 2003. Multiple UAV cooperative search under collision avoidance and limited range communication constraints," *Proceedings of the IEEE Conference on Decision and Control*.

[14] Vincent, P., Rubin, I. , 2004. A Framework and Analysis for Cooperative Search Using UAV Swarms. *ACM Symposium on Applied Computing.*

[15] Finn, A. et al., 2007. Design Challenges for an Autonomous Cooperative of UAVs, In *Proceedings of the International Conference on Information, Decision and Control*.

[16] Liao, Y., Jin, Y., Minai, A. and Polycarpou, M. 2005. Information Sharing in Cooperative Unmanned Aerial Vehicle Teams, *Decision and Control.*

[17] Velagapudi, Prokopiev, Sycara, Scerri. Maintaining Shared Belief in a Large Mutiagent Team, Fusion 2007. (see Scerri: http://www.cs.cmu.edu/~pscerri/Publications.html)

[18] Sycara, K. et al.2007. An Analysis and Design Methodology for Belief Sharing in Large Groups, *Fusion 2007*, Quebec.

[19] Intanagonwiwat, C. et al. 2003. Directed diffusion for wireless sensor networking. *IEEE/ACM Trans. on Networking*, 11(1):2–16.

[20] Xuan, P., Lesser, V., Zilberstein, S., 2001. Communication decisions in multi-agent cooperation: Model and experiments. In *Proceedings of the Fifth International Conference on Autonomous Agents*, Montreal, Canada.

[21] Brooks, A., Makarenko, A., Williams, S., Durrant-Whyte, H. 2006. Parametric POMDPs for Planning in Continuous State Spaces In *Robotics and Autonomous Systems*, 54(11).

[22] Bernstein, D., Givan, R., Immerman, N., Zilberstein, S. 2002. The complexity of decentralized control of Markov decision processes. *Mathematics of Operations Research,* 27 (4), 819-840.

[23] Polycarpou, M., Yang, Y., and Passino, K., 2001. A cooperative search framework for distributed agents," *IEEE Int. Symp. Intelligent Control*, Mexico City, Mexico.

[24] Yang, Y., Minai, A., Polycarpou, M. 2005. Evidential Map Building Approaches for Multi-UAV Cooperative Search. *Proceedings of the American Control Conference*.

[25] Yang, Y., Minai, A., and Polycarpou, M. 2004. Decentralized Cooperative Search by Networked UAVs in an Uncertain Environment, *American Control Conference*.

[26] Yang, Y., Polycarpou, M., Minai, A. 2007. Multi-UAV Cooperative Search Using an Opportunistic Learning Method. *Journal of Dynamic Systems, Measurement, and Control,* 129(5).

[27] Cover, T. and Thomas, J., 2006. *Elements of Information-Theory*, 2nd edition, Wiley.

[28] Potvin, J.-Y. and Bengio, S., 1996. The Vehicle Routing Problem with Time Windows Part II: Genetic Search, *INFORMS Journal on Computing* 8, 165–172.

[29] Goldberg, D., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning.* New York: Addison-Wesley.