TRUSTED AI SE

SERC/AIRC Phase II

S. Cornejo, Z. Alizadeh, A. Youssef, C.J. Buss, J. Gregory
A. Jalilzadeh, A. Salado, P. Satam
Department of Systems and Industrial Engineering

THE UNIVERSITY OF ARIZONA

PHASE II – GOAL

Design of the decision system; Human-machine teaming; Resilience

ASSUMPTIONS

UGV is a scarce resource
UAV is a fast, multi-spectral video collection system
AI performance data corresponds to UAV
Human SME reviews video from UAV
Human SME gets feedback from UGV

APPROACH

Mission Engineering in a DE environment
Multidisciplinary team

THE UNIVERSITY OF ARIZONA

MISSION NEEDS & REQUIREMENTS

MN-1 A battalion needs to safely traverse between two points as quickly as possible.

ID	Mission Requirement	Justification
MR-1	**Time to clear a path** The mission needs to declare a path as clear for a battalion to move from point A to point B in less than TBD h. *Note 1*: Less time is preferred to more time. *Note 2*: Points A and point B are inputs to the mission.	Identified by the sponsor in the documentation provided to the team.
MR-2	**Effectiveness of path clearance** The mission needs to clear a path for a battalion to move from point A to point B with likelihood over TBD%. *Note 1*: Likelihood refers to a mine being left uncleared on the path. *Note 2*: More confidence is preferred to less confidence. *Note 3*: Points A and point B are inputs to the mission.	Not explicitly identified by the sponsor but derives from the need to safely traverse the path. Clearly, declaring a path as cleared without being so would be inadequate.
MR-3	**Soldier trustworthiness*** The mission needs to yield trustworthiness above TBD to the soldiers that are to traverse the path. *Note*: Higher trustworthiness is preferred to lower trustworthiness.	Not explicitly identified by the sponsor but implicit in the documentation. From a mission perspective, success in the soldiers traversing the path will depend on the extent to which they trust the path has been cleared.

TERRAIN

Type
AI confidence
Human confidence
Mine/No mine

Confidence mean
Confidence variance

MOEs

0: Time to traverse path
1: Time to clear path
2: Clearance effectiveness
3: Soldier trust

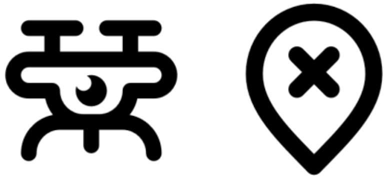
THE UNIVERSITY OF ARIZONA

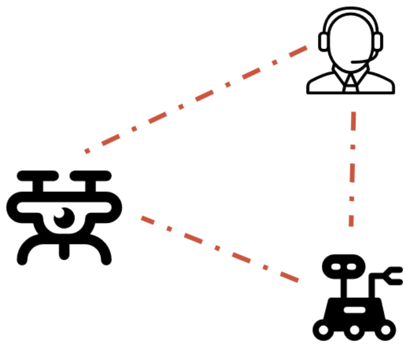
MISSION THREADS

Grassy, wooded, swampy, rocky…
Snowy!

MT1. Only nodes with high performance by the human.
MT2. Only nodes with high performance by the AI.
MT3. First half good performance human, second halg good performance AI.
MT4. Only nodes with high performance by both.
MT5. Nodes with poor performance by both.
MT6. Nodes with no characterized performance of any of them.
MT7. Random location of nodes (per performance of AI and human, mixed performance).

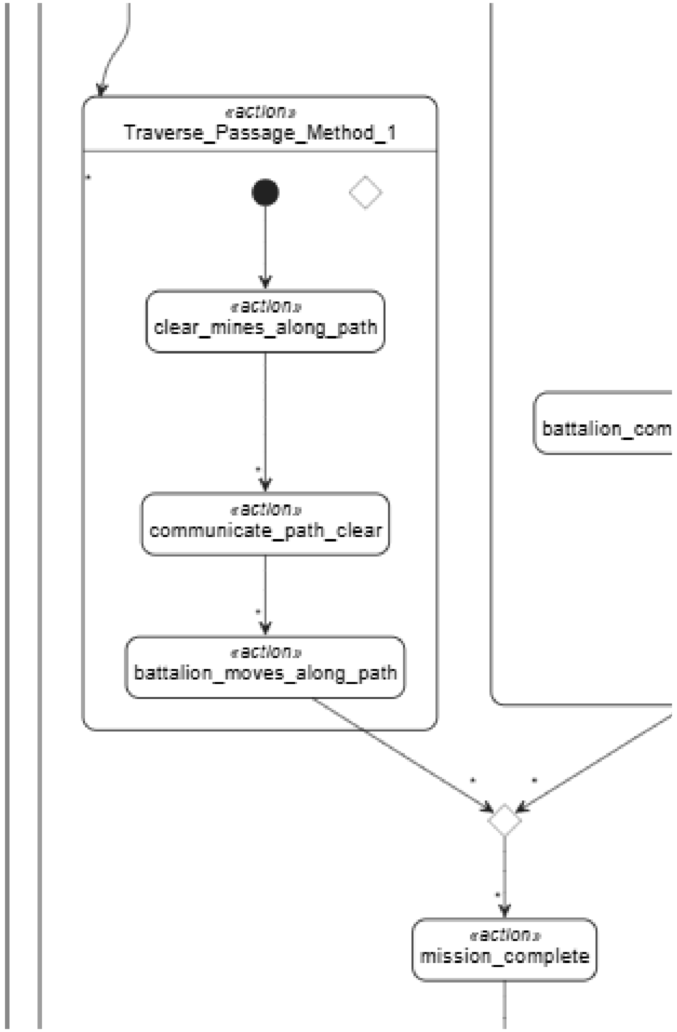
Security/compromised actors

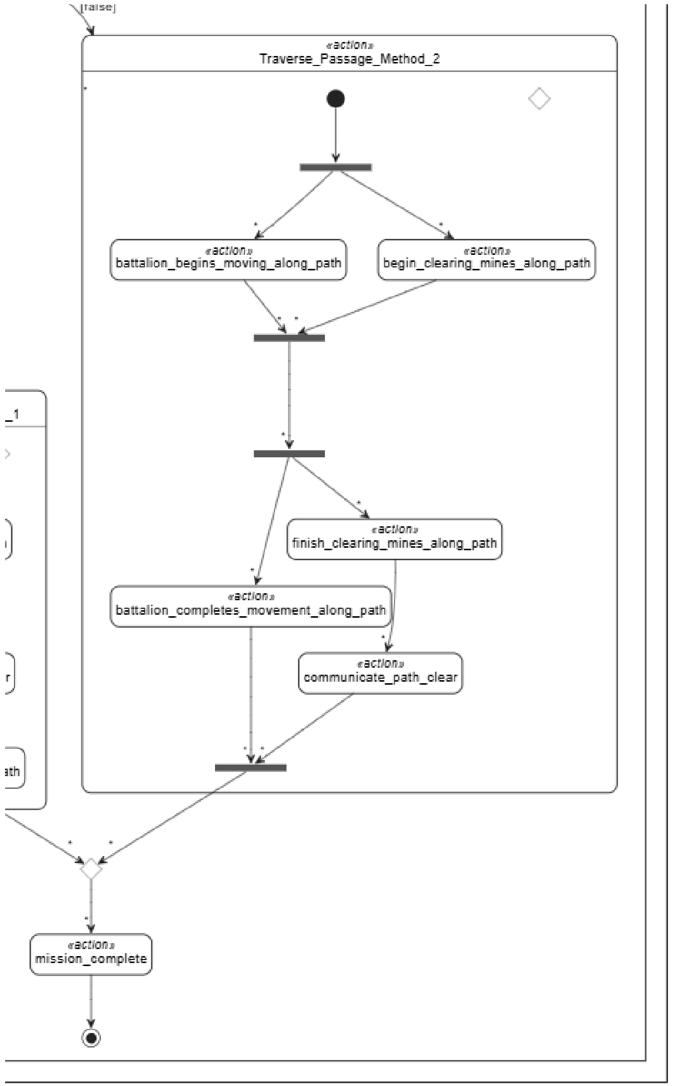




THE UNIVERSITY OF ARIZONA

CONOPS


«action»
Traverse_Passage_Method_1
«action»
clear_mines_along_path
«action»
communicate_path_clear
«action»
battalion_moves_along_path
battalion_com
«action»
mission_complete


«action»
Traverse_Passage_Method_2
«action»
battalion_begins_moving_along_path
«action»
begin_clearing_mines_along_path
«action»
finish_clearing_mines_along_path
«action»
battalion_completes_movement_along_path
«action»
communicate_path_clear
«action»
mission_complete

MISSION MODEL

$$\mathrm{MOE0} = t_n^B$$

$$s.t.: t_j^{UGV} = t_{ij}^{UGV} x_{ij}^{UGV,(1)} + t_{fm} x_{ij}^{UGV,(1)} + t_{cl} x_{ij}^{UGV,(2)} + \max\left(tp_j^{UAV}, tp_j^{huma}\right.$$

$$t_j^{UGV} = t_{ij}^{UGV} x_{ij}^{UGV,(1)} + \max\left(tp_j^{UAV}, tp_j^{human}, t_i^{UGV}\right), \forall\,(i,j) \in$$

$$t_j^B = \max(\max(t_j^{UGV}, t_i^B) + t_{ij}^B x_{ij}^B +, d_j^B c_j^B c_j^{UGV}), \forall\,(i,j) \in$$

$$c_i^{UGV} = \begin{cases} \dfrac{p_i^{UAV}}{per_i^{UAV}} \dfrac{p_i^{human}}{per_i^{human}}, x_i^{UGV} = 1, p_i^{UAV} < \epsilon^{mine} \vee p_i^{human} < \epsilon \\ 1 - m^{UGV}, x_i^{UGV} = 1, p_i^{UAV} \geq \epsilon^{mine} \vee p_i^{human} \geq \epsilon^{mine} \end{cases}$$

$$p_i^{actor} = \begin{cases} \mu_i^{actor} - \sigma_i^{actor^2}, \mu_i^{actor} \geq 0.5, h_i^{actor} = 1 \\ \mu_i^{actor} + \sigma_i^{actor^2}, \mu_i^{actor} < 0.5, h_i^{actor} = 1, \forall\, actor \in \{huma \\ 0.5, h_i^{actor} = 0 \end{cases}$$

$$d_i^B = \begin{cases} 0, i \in V \backslash \bar{V} \\ \mathcal{U}(d_{min}, \infty), i \in \bar{V} \end{cases}$$

$$tp_i^{UAV} = (t_i^{UAV} + t_{pred}^{UAV}) h_i^{UAV}, i \in V$$

$$tp_i^{human} = (t_i^{UAV} + t_{pred}^{human}) h_i^{human}, i \in V$$

$$t_{i+1}^{UAV} = t_i^{UAV} + t_{ij}^{UAV} x_{ij}^{UAV}, \forall\,(i,j) \in E$$

$$MOE1 = t_n^B$$

$$s.t.: t_j^{UGV} = t_{ij}^{UGV} x_{ij}^{UGV,(1)} + t_{fm} x_{ij}^{UGV,(1)} + t_{cl} x_{ij}^{UGV,(2)} + \max\left(tp_j^{UAV}, tp_j^{human}, t_i^{UGV}\right), \forall\,(i,j) \in \bar{E}$$

$$t_j^{UGV} = t_{ij}^{UGV} x_{ij}^{UGV,(1)} + \max\left(tp_j^{UAV}, tp_j^{human}, t_i^{UGV}\right), \forall\,(i,j) \in E \backslash \bar{E}$$

$$\mathrm{MOE2} = \sum_{i \in V} c_i^{UGV} x_i^{UGV}$$

$$c_i^{UGV} = \begin{cases} \dfrac{p_i^{UAV}}{per_i^{UAV}} \dfrac{p_i^{human}}{per_i^{human}}, x_i^{UGV} = 1, p_i^{UAV} < \epsilon^{mine} \vee p_i^{human} < \epsilon^{mine} \\ 1 - m^{UGV}, x_i^{UGV} = 1, p_i^{UAV} \geq \epsilon^{mine} \vee p_i^{human} \geq \epsilon^{mine} \end{cases}, \forall i \in V$$

$$\mathrm{MOE3} = \sum_{i \in V} c_i^B x_i^B$$

$$FN^{actor} = \sum_{i \in v'} \mathbb{I}_{(p_i^{actor} < e^{mine})} \quad , \quad FP^{actor} = \sum_{i \in v \backslash v'} \mathbb{I}_{(p_i^{human} > e^{mine})}, \forall\, actor \in \{UAV, human\}$$

$$c_i^B = f((FP^{UAV}, FN^{UAV}), (FP^{human}, FN^{human}))$$

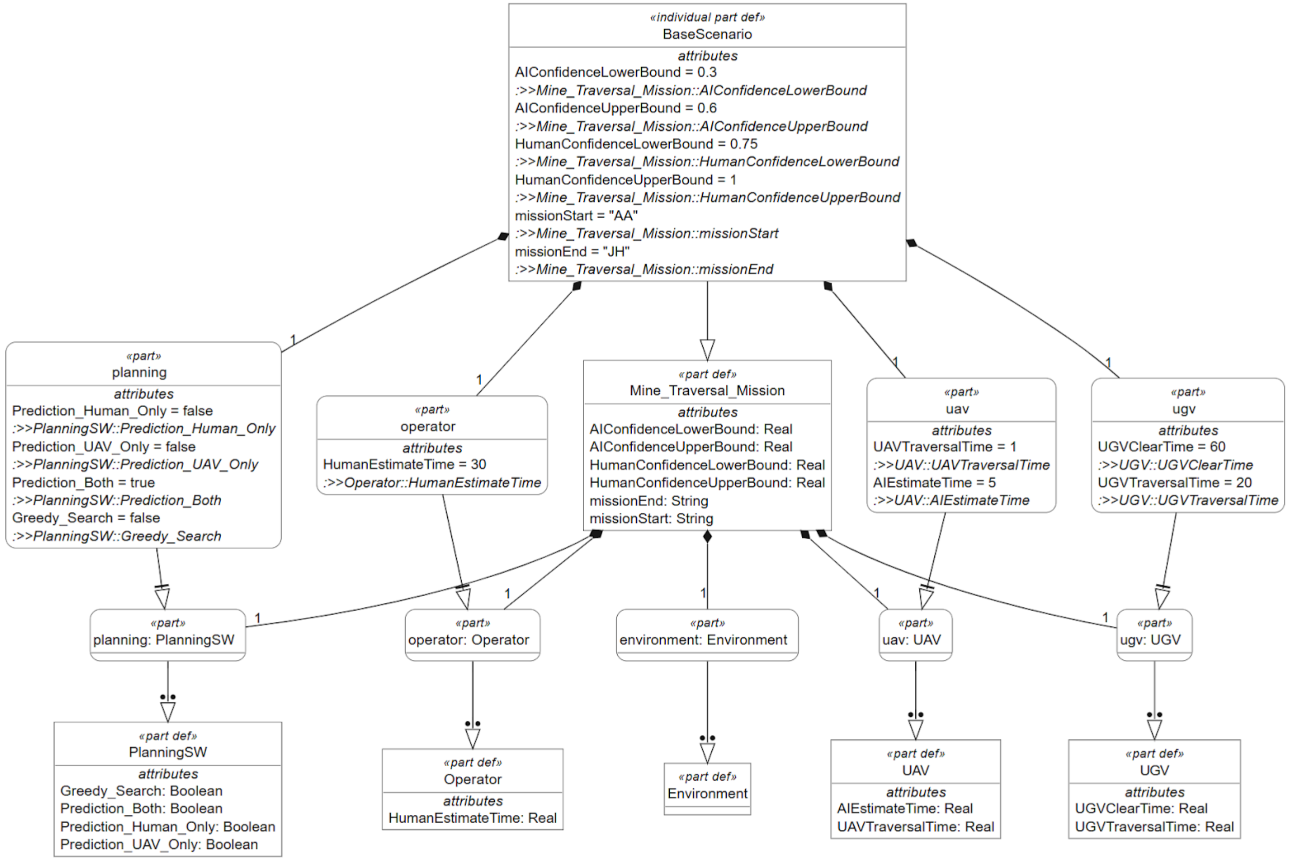
THE UNIVERSITY OF ARIZONA

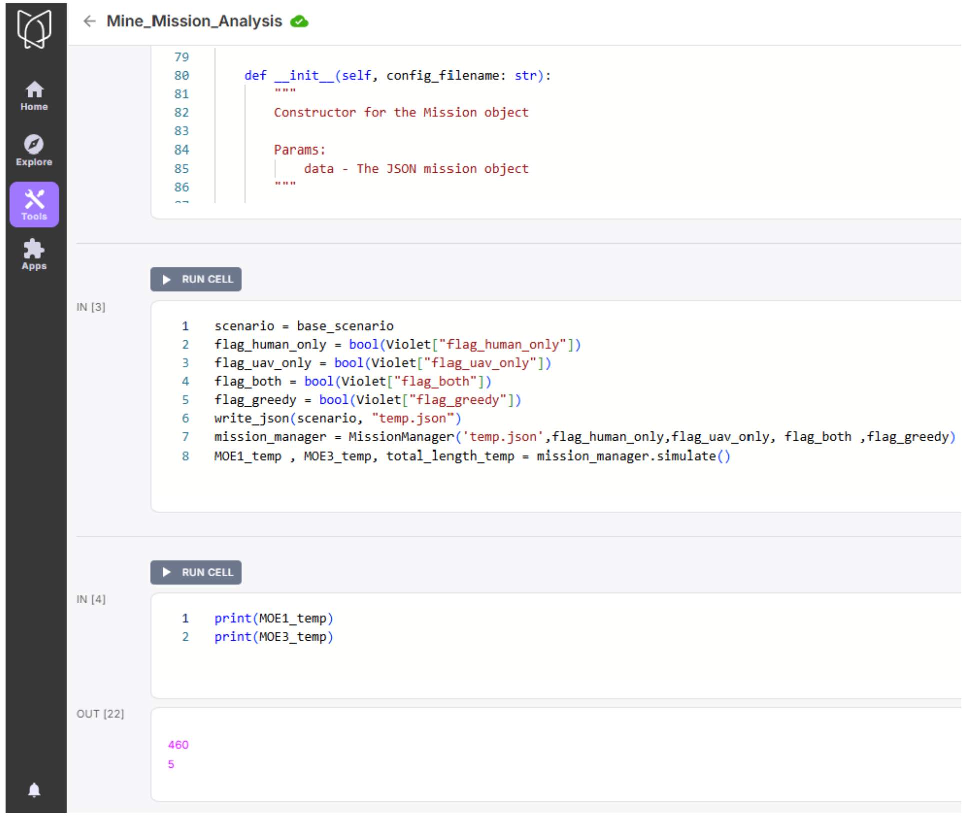
ARCHITECTURE ALTERNATIVES

Allocation of prediction tasks	*Full AI, Full Human, Dynamic*
Task sequencing and parallelization	*Serial, Parallel*
Centralization of task coordination	*Centralized, Decentralized*
Prioritization of prediction tasks	*From A to B, Unconstrained*

THE UNIVERSITY OF ARIZONA

SIMULATION & ANALYSIS



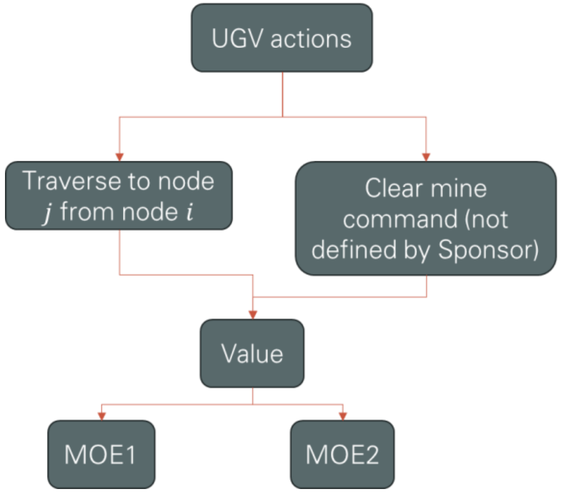


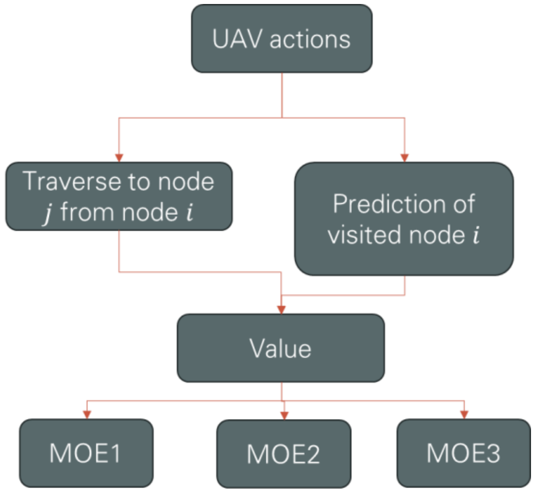
SIMULATION & ANALYSIS

Scenarios = Mission threads X Architectures

Monte Carlo – 100 runs for each scenario

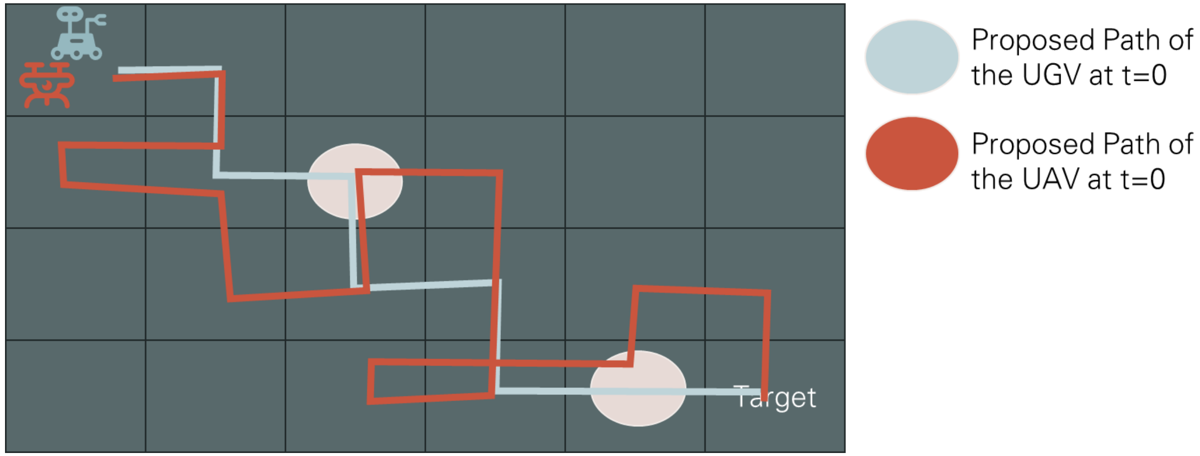
min(MOE0), min(MOE1), max(MOE2), max(MOE3)

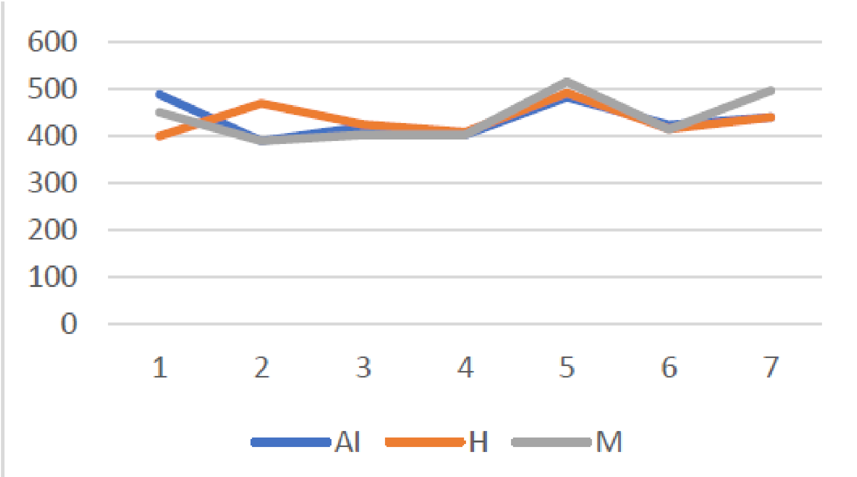


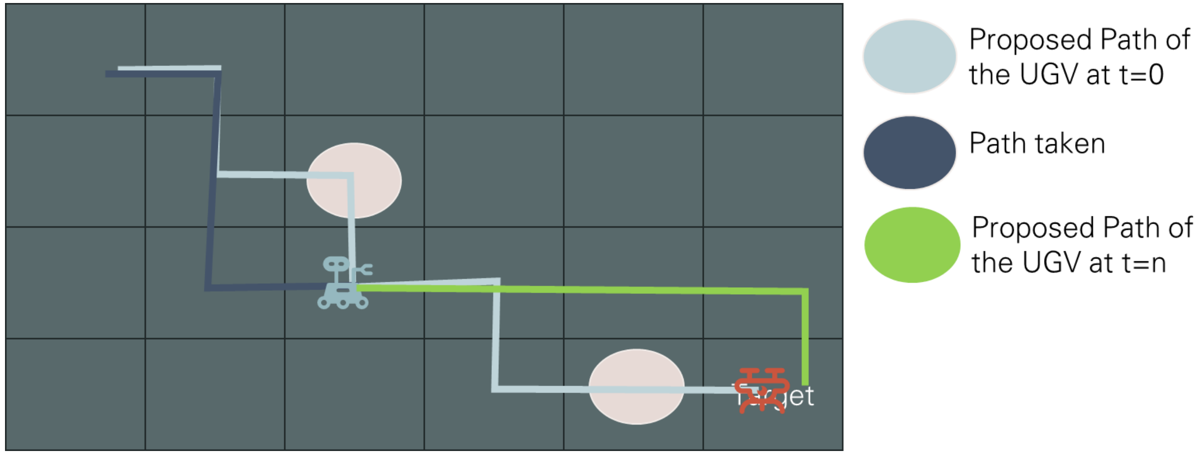


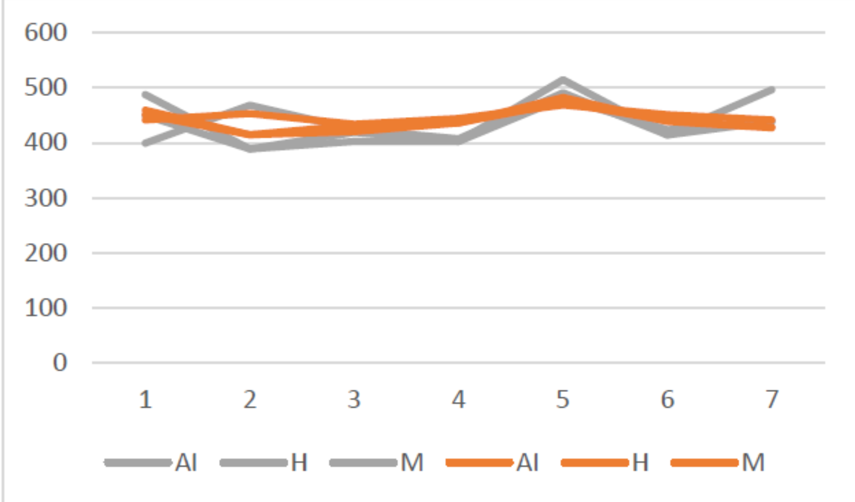
	Architectures					
	Full AI Not Greedy	Full Human Not Greedy	Intermediate Not Greedy	Full AI Greedy	Full Human Greedy	Intermediate Greedy
Mission Thread 1 Optimal: 282	MOE0: TBD MOE1: 487.6 MOE2: TBD MOE3: 5.1	MOE0: TBD MOE1: 399.6 MOE2: TBD MOE3: 0.48	MOE0: TBD MOE1: 449.6 MOE2: TBD MOE3: 2.33	MOE0: TBD MOE1: 459.6 MOE2: TBD MOE3: 5.01	MOE0: TBD MOE1: 441.8 MOE2: TBD MOE3: 4.63	MOE0: TBD MOE1: 451.2 MOE2: TBD MOE3: 4.15
Mission Thread 2 Optimal: 284.8	MOE0: TBD MOE1: 389 MOE2: TBD MOE3: 0.57	MOE0: TBD MOE1: 468.6 MOE2: TBD MOE3: 4.89	MOE0: TBD MOE1: 389 MOE2: TBD MOE3: 0.57	MOE0: TBD MOE1: 414.8 MOE2: TBD MOE3: 3.02	MOE0: TBD MOE1: 453.4 MOE2: TBD MOE3: 5.13	MOE0: TBD MOE1: 414.8 MOE2: TBD MOE3: 3.02
Mission Thread 3 Optimal: 279.8	MOE0: TBD MOE1: 419.2 MOE2: TBD MOE3: 2.53	MOE0: TBD MOE1: 424.4 MOE2: TBD MOE3: 2.66	MOE0: TBD MOE1: 402.4 MOE2: TBD MOE3: 1.28	MOE0: TBD MOE1: 432 MOE2: TBD MOE3: 4.33	MOE0: TBD MOE1: 434 MOE2: TBD MOE3: 4.67	MOE0: TBD MOE1: 420.6 MOE2: TBD MOE3: 3.77
Mission Thread 4 Optimal: 280.4	MOE0: TBD MOE1: 402.8 MOE2: TBD MOE3: 0.46	MOE0: TBD MOE1: 407.8 MOE2: TBD MOE3: 0.86	MOE0: TBD MOE1: 402.8 MOE2: TBD MOE3: 0.46	MOE0: TBD MOE1: 436.0 MOE2: TBD MOE3: 3.79	MOE0: TBD MOE1: 444.2 MOE2: TBD MOE3: 4.93	MOE0: TBD MOE1: 436 MOE2: TBD MOE3: 3.79
Mission Thread 5 Optimal: 288.4	MOE0: TBD MOE1: 481.6 MOE2: TBD MOE3: 4.85	MOE0: TBD MOE1: 491 MOE2: TBD MOE3: 5.6	MOE0: TBD MOE1: 514.8 MOE2: TBD MOE3: 5.26	MOE0: TBD MOE1: 473 MOE2: TBD MOE3: 5.34	MOE0: TBD MOE1: 468.8 MOE2: TBD MOE3: 5.47	MOE0: TBD MOE1: 482.6 MOE2: TBD MOE3: 5.3
Mission Thread 6 Optimal: 288.6	MOE0: TBD MOE1: 423.6 MOE2: TBD MOE3: 2.65	MOE0: TBD MOE1: 414.4 MOE2: TBD MOE3: 2.62	MOE0: TBD MOE1: 414.2 MOE2: TBD MOE3: 1.52	MOE0: TBD MOE1: 440.4 MOE2: TBD MOE3: 3.89	MOE0: TBD MOE1: 451.6 MOE2: TBD MOE3: 4.86	MOE0: TBD MOE1: 439 MOE2: TBD MOE3: 3.56
Mission Thread 7 Optimal: 285	MOE0: TBD MOE1: 439 MOE2: TBD MOE3: 2.31	MOE0: TBD MOE1: 430 MOE2: TBD MOE3: 2.51	MOE0: TBD MOE1: 406.4 MOE2: TBD MOE3: 0.97	MOE0: TBD MOE1: 429 MOE2: TBD MOE3: 3.6	MOE0: TBD MOE1: 442 MOE2: TBD MOE3: 4.82	MOE0: TBD MOE1: 427.2 MOE2: TBD MOE3: 3.38

EXAMPLE OF SOLUTION STRATEGIES & ASSESSMENT









THE UNIVERSITY OF ARIZONA

NEXT

Complete analysis and assessment

Increase focus on compromised scenarios

Study engineering teaming

THANK YOU

THE UNIVERSITY OF ARIZONA