

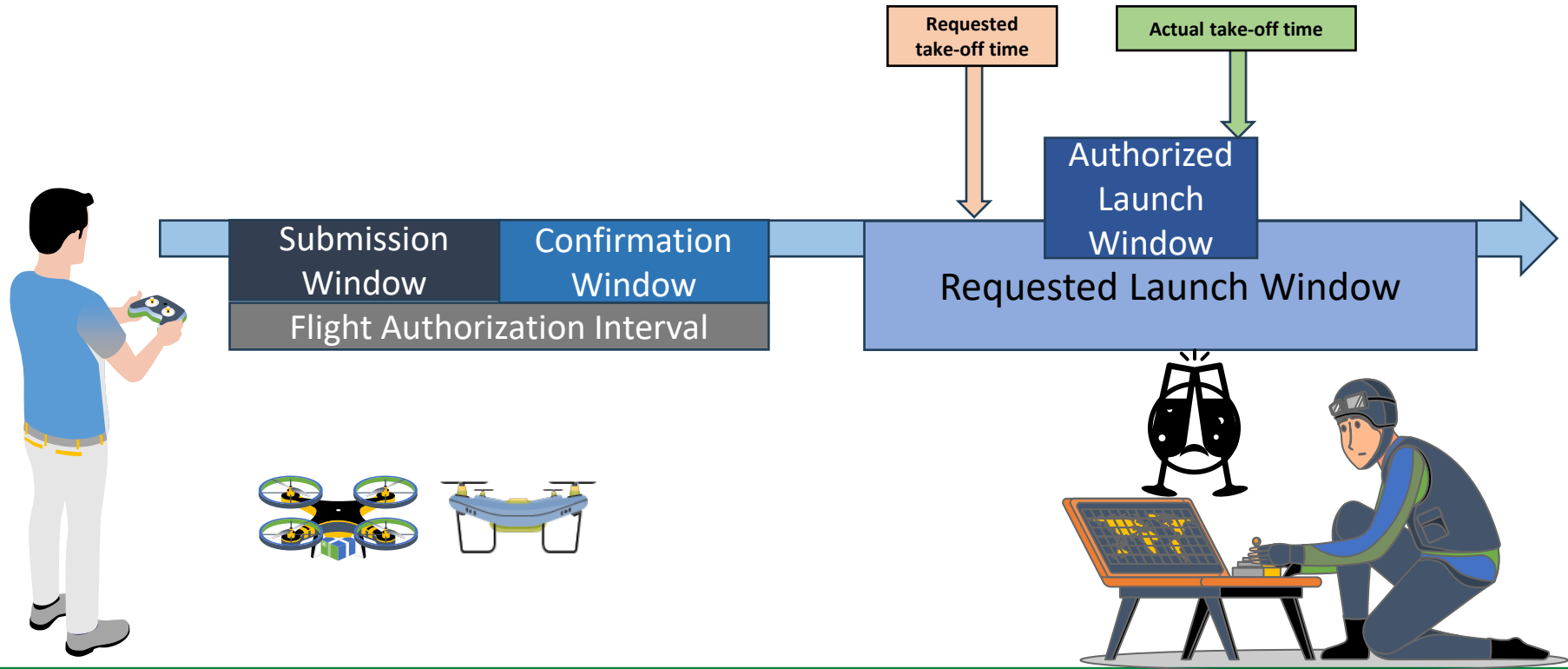
From Rejection to Re-Accommodation: A Generate-and-Negotiate Framework for Conflict-Free U-space

Zhiqiang Liu

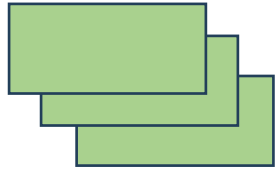
Juan Jose Ramos Gonzalez

Jose Luis Munoz-Gamarra

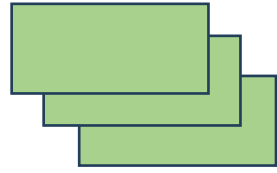
Strategic Planning & Deconflicting Service



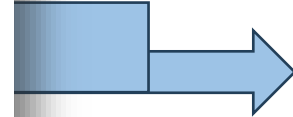
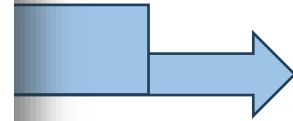
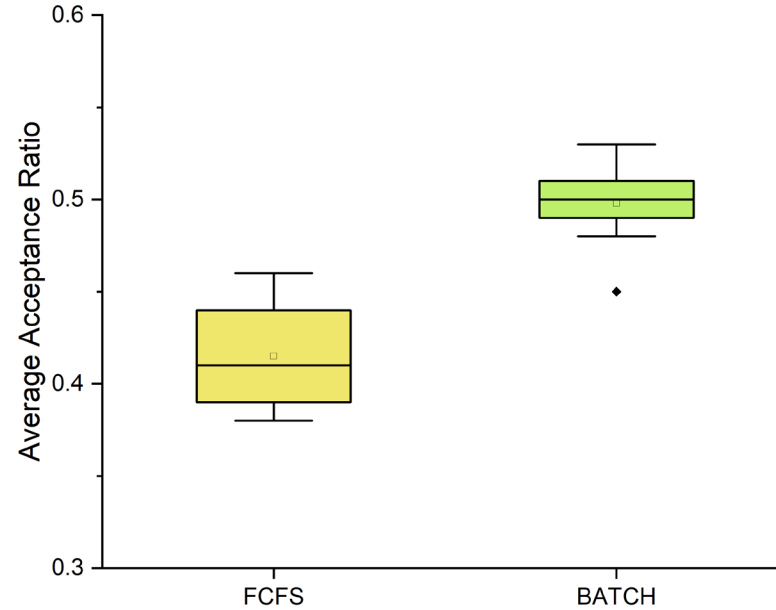
FCFS VS BATCH



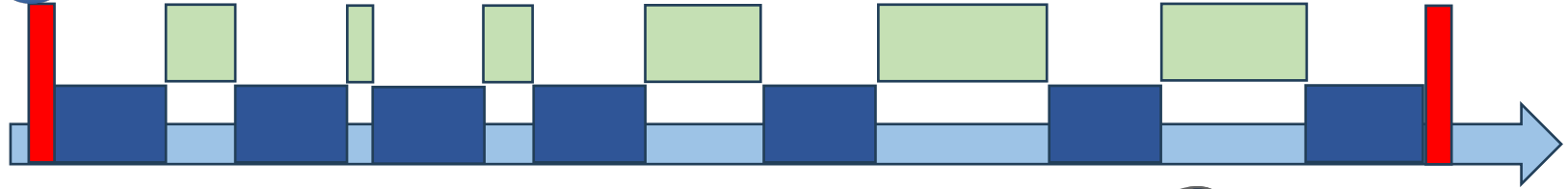
First Come First Served policy
Missions are subject to approval



Batch submission:
A set of missions is in the p



Possible Airspace Usage

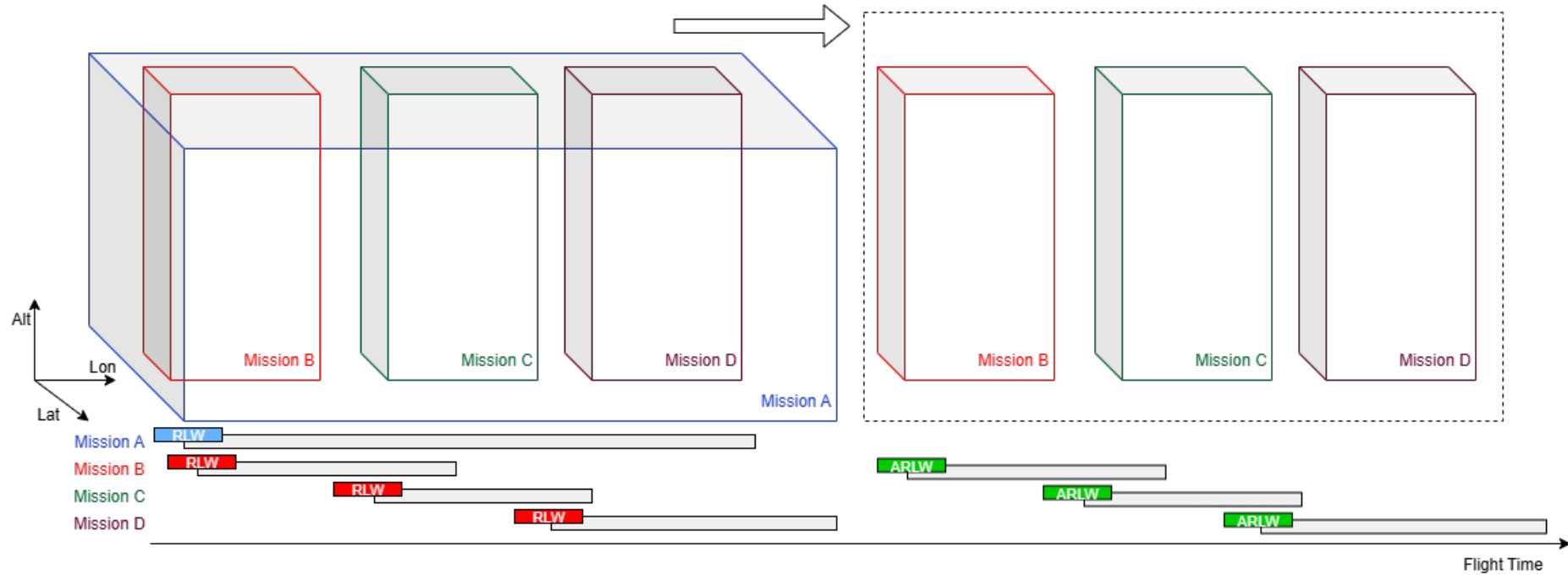


Occupied
Airspace

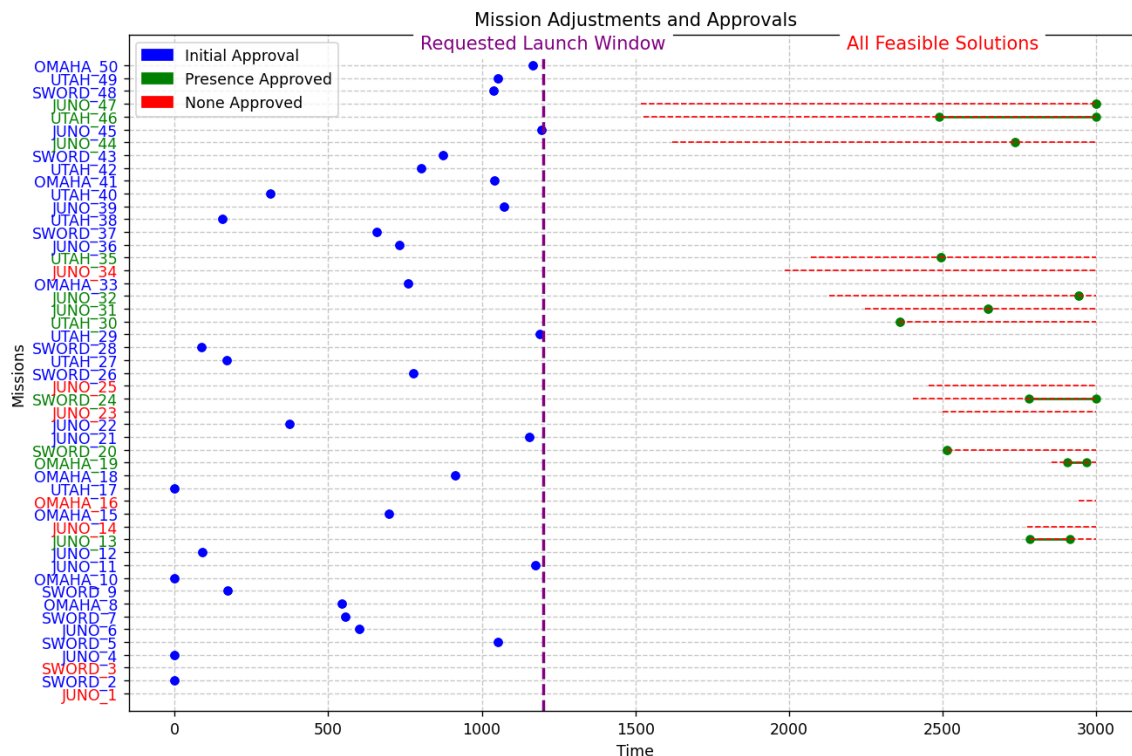
Free
Airspace



Alternative Requested Launch Window



Single Interval Alternative Solving Model



50 missions in total

Initial 32/50 (64%) approved

Alternative found 11 within 18 RFPs

At the end 43/50 (86%) approved

22% increase



Global Solution Table

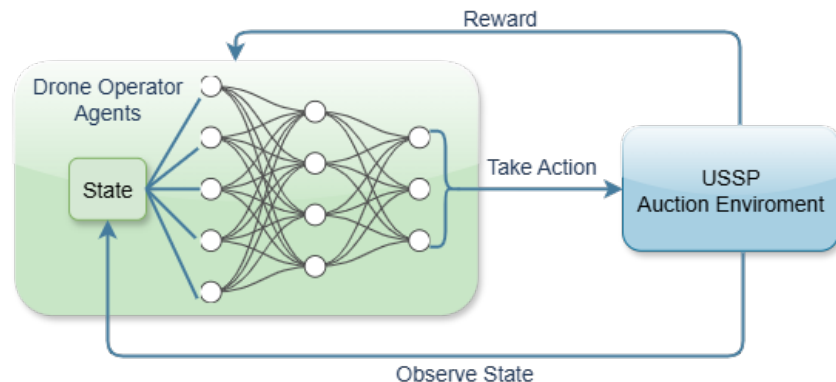
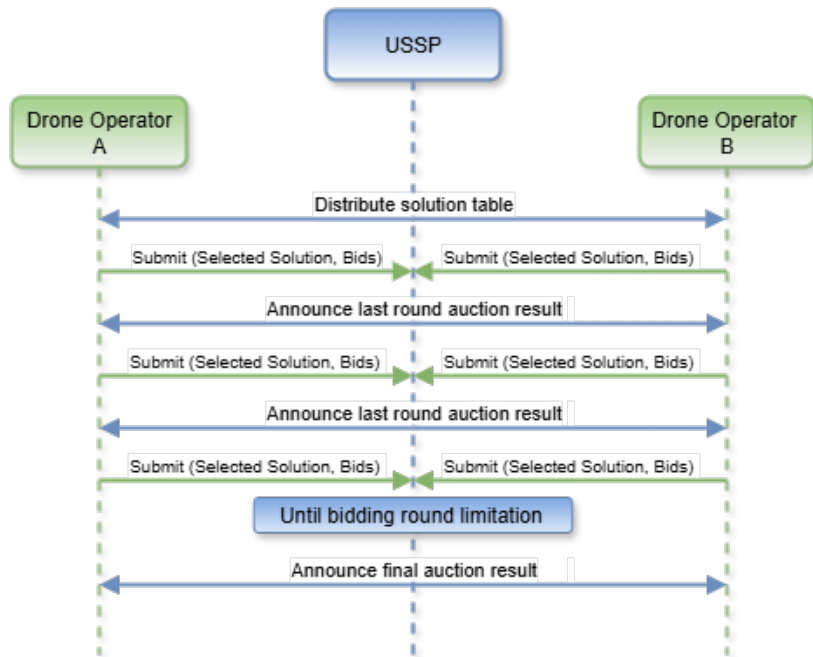


Drone Operators (DOs)	DO-A		DO-B	
Rejected Flight Plans (RFPs)	RFP 1	RFP 2	RFP 3	RFP 4
Solution 1	ARLW 1,1	/	ARLW 1,3	ARLW 1,4
Solution 2	/	ARLW 2,2	/	/
Solution 3	/	ARLW 3,2	ARLW 3,3	ARLW 3,4
Solution 4	ARLW 4,1	ARLW 4,2	ARLW 4,3	/
Solution 5	ARLW 5,1	/	ARLW 5,3	/

Alternative Requested Launch Window (ARLW):
Based on original Requested Launch Window, [-50 mins, +50 mins]



Negotiation Framework



P-coins is used to evaluate the bid of agents.

DO Agent Type

Achiever

- Prioritises mission success over bidding cost (risk-seeking operator).
- Bids aggressively to win as many ARLWs as possible.

Optimizer

- Seeks a balance between mission acceptance and P-coin expenditure.
- Chooses solutions with good re-accommodation capacity at reasonable bid levels.

Economizer

- Highly cost-conscious, bids conservatively to preserve P-coins.
- Enters competition only when the expected gain clearly justifies the cost.

Normalizer

- Neutral “baseline” operator with no strong bias toward winning or saving.
- Uses stable, mid-range bids and does not actively search for competitive advantage.

Reward

Round reward: Immediate reward based on chosen solution in each round.

$$R_{\text{round}}^i(j) = \begin{cases} 0.01 + \frac{M_{\text{fly}}(j)}{M_{\text{reject}}(j)} * 0.01 & \text{if } M_{\text{max}}(j) = M_{\text{max_global}}, \\ 0.005 + \frac{M_{\text{fly}}(j)}{M_{\text{reject}}(j)} * 0.005 & \text{if } M_{\text{max}}(j) \geq 0.8 * M_{\text{max_global}}, \\ -0.01 & \text{otherwise.} \end{cases}$$

Where:

- $M_{\text{fly}}(j)$ number of RFPs that could be re-accommodate under the selected solution.
- $M_{\text{reject}}(j)$ number of RFPs in the selected solution.
- $M_{\text{max}}(j)$ maximum RFPs that could be re-accommodate in this solution.

Reward

Individual reward: Agent-specific reward based on its role at the end of the negotiation process.

$$R_{\text{individual}}^i = (w_1 R_{\text{cost_penalty}} + w_2 R_{\text{bonus}} + w_3 R_{\text{lose_penalty}}) / 10,$$

where:

- $R_{\text{cost_penalty}}$ for excessive spending of P-coins. Calculate as $-(\text{final bid}/5)$ if DO is the winner.
- R_{bonus} winning bonus for securing solution. Calculate as $+10$ if DO wins the auction.
- $R_{\text{lose_penalty}}$: if DO lose the auction will receive penalty. Calculate as -10 if DO lose the auction.

Agent role	Cost weight (w_1)	Bonus weight (w_2)	Lose penalty weight (w_3)
Optimizer	0.5	1.0	1.0
Achiever	0.0	1.5	1.0
Economizer	1.5	0.5	1.0
Normalizer	0.5	0.5	0.5

Reward

Global reward: encourages fair competition and efficient resource allocation, will be given at the end of the negotiation.

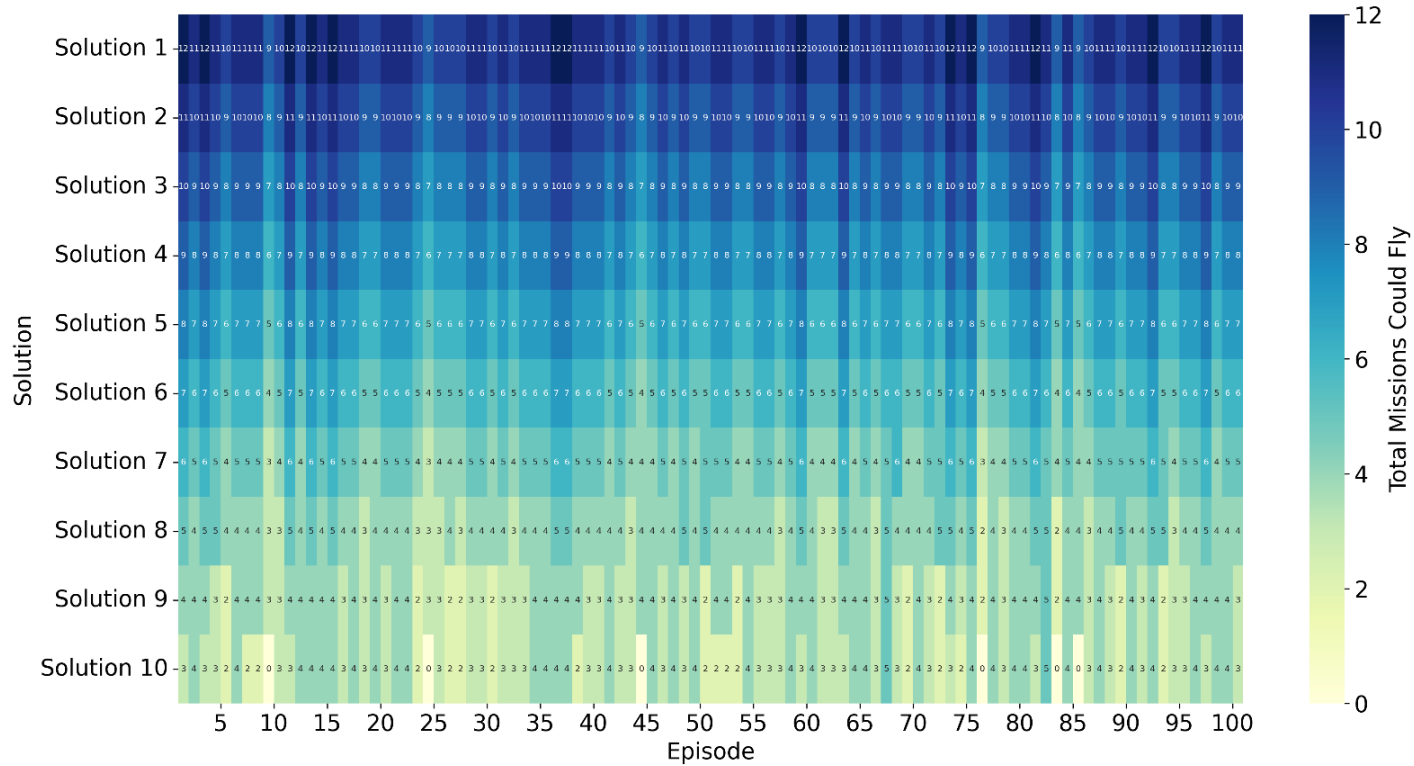
$$R_{\text{global}}^i = \begin{cases} \frac{M_{\text{total}}}{M_{\text{max_global}}}, & \text{if } M_{\text{total}} \geq 0.8 * M_{\text{max_global}}, \\ -(1 - \frac{M_{\text{total}}}{M_{\text{max_global}}}), & \text{otherwise.} \end{cases}$$

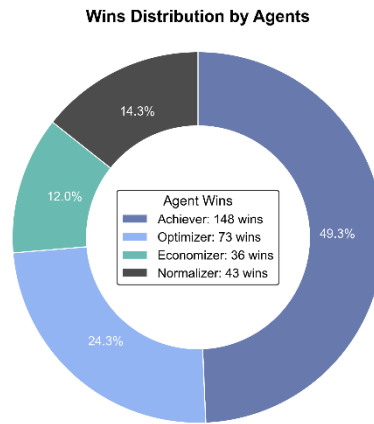
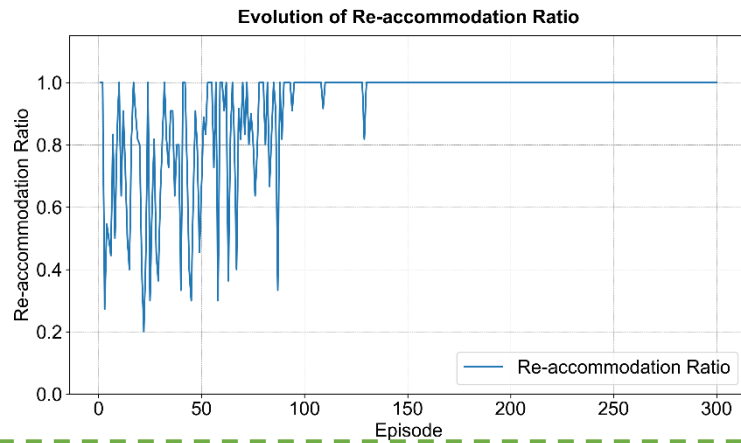
Where:

- M_{total} the total number of successfully re-accommodated RFPs across all DOs.
- $M_{\text{max_global}}$: the global maximum possible RFPs that could be re-accommodate.
- $0.8 * M_{\text{max_global}}$: a threshold ensuring that at least 80% of the total RFPs capacity is utilized before positive rewards are granted.

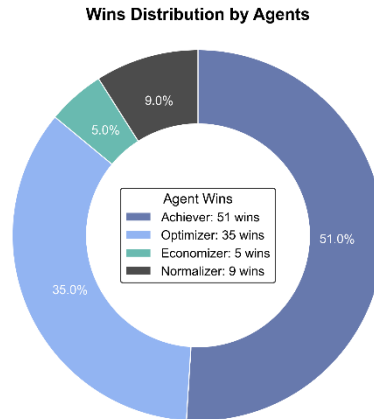
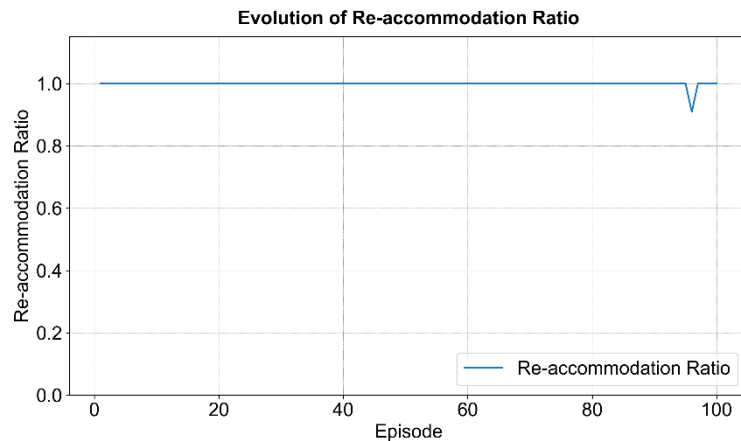
Global Solution Table

Global Solution per Episode

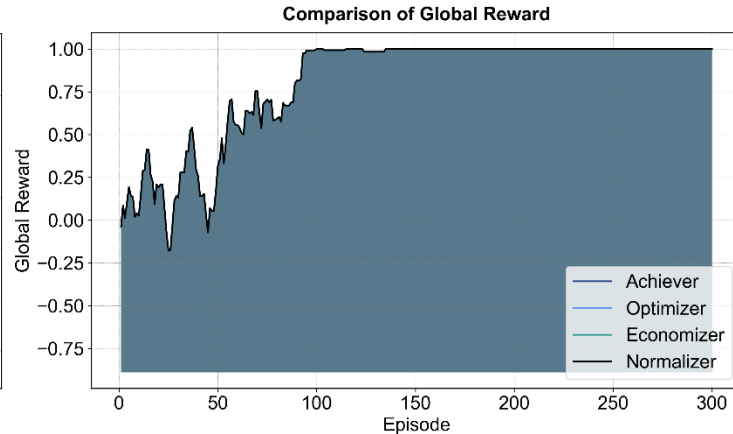
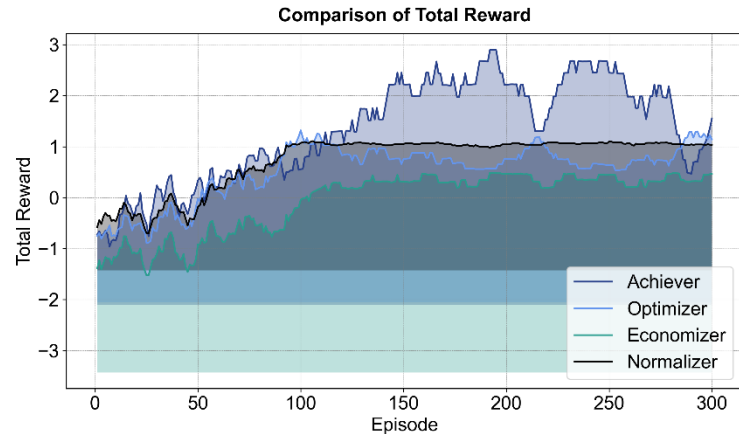




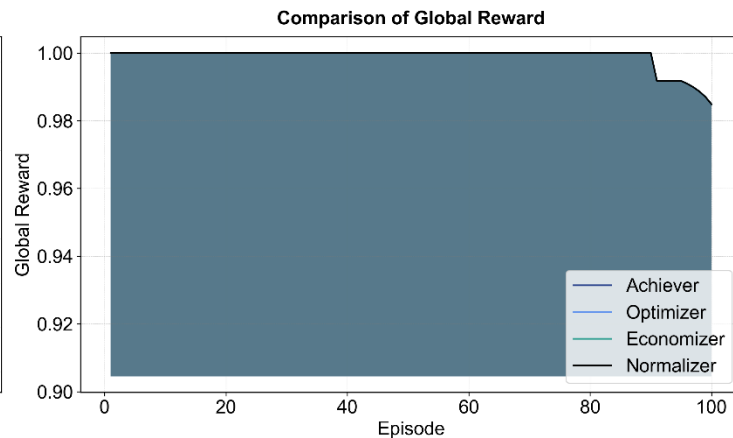
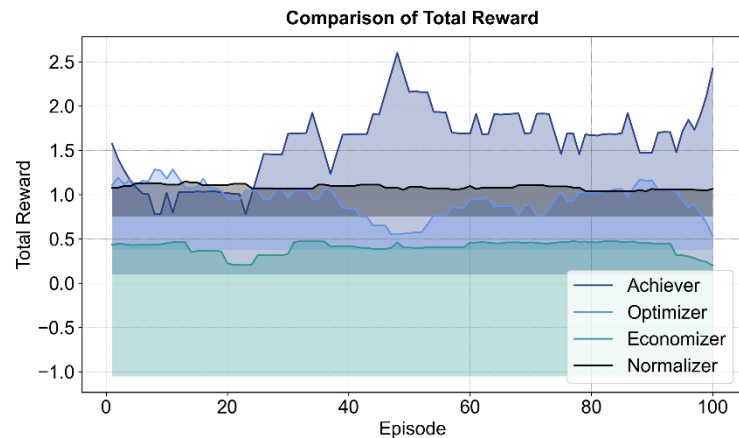
Training



Validation



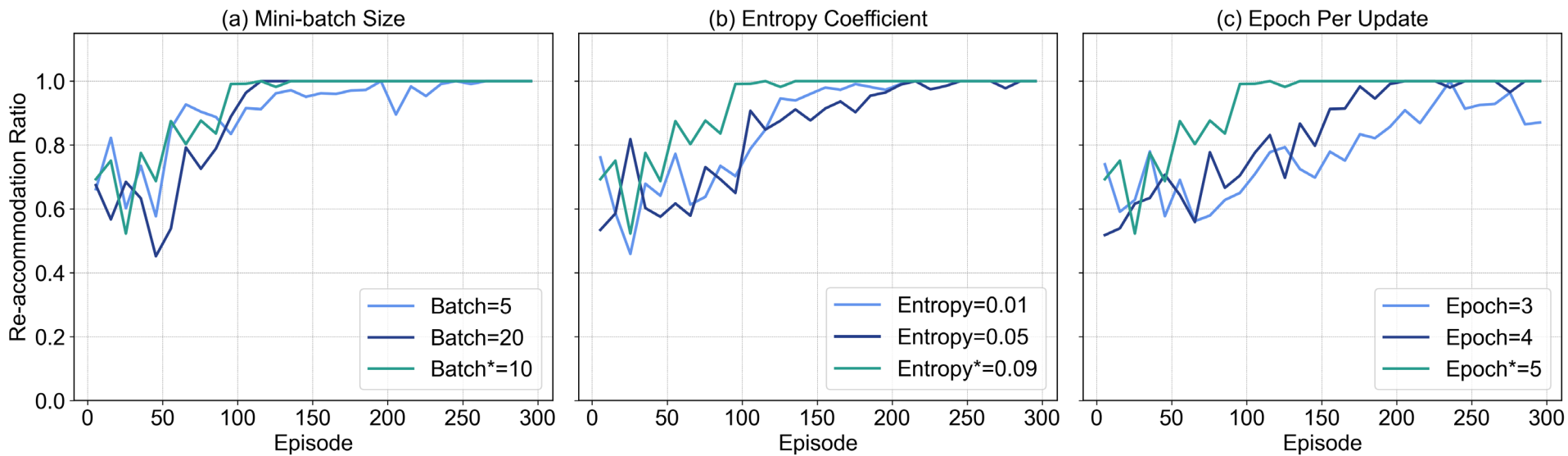
Training



Validation

Sensitivity

Sensitivity of Re-accommodation Ratio to PPO Training Settings



Conclusions

- The Single-Interval model generates conflict-free ARLWs by time-shifting along the same route, providing the solution tables for the rejected missions while keeping computation sub-second.
- On top of solution tables, the DRL-MANF framework models different DO as heterogeneous agents and uses a PPO-based sealed-bid auction to learn effective bidding strategies.
- In simulations, the “Generate and Negotiate” pipeline reduces rejections and re-accommodate previously rejected flights without violating separation or capacity constraints.
- Future work will refine the P-coin mechanism and train agents with alternative learning algorithms to compare efficiency, stability, and fairness.

Thank you for your attention!

Zhiqiang Liu

Email: zhiqiang.liu@uab.cat

LinkedIn: Zhiqiang Liu

