Evolved AI for Model-Based Reinforcement Learning

Randal Allen
Lone Star Analysis
Orlando, FL
rallen@lone-star.com

Zachry Engel, Eric Haney
Lone Star Analysis
Addison, TX
zengel@lone-star.com, ehaney@lone-star.com

ABSTRACT

Challenges in aviation maintenance include "just in time" maintenance and logistic support chain requirements. Optimizing the correct level of investment in repairable assets while minimizing storage footprint requirements at multiple sites is highly desirable. Particularly challenging is accounting for the number of spare assets needed when there is a lack of availability due to shortages in the industrial supply chain, an inaccurate forecast, and/or uncertain maintenance turn-around times. This multi-dimensional, time-varying problem is currently addressed with human maintenance schedulers and heuristics – leading to situations with ambiguous actions and unintended long-term consequences.

This paper demonstrates how model-based reinforcement learning is used to choose the optimal aviation maintenance policy when considering both immediate and subsequent costs. Aircraft states are defined according to equipment operational capability (EOC) codes which classify the degradation of equipment mission capability. Based on the EOC codes, actions are selected from a set of possible choices. The consequence of an action is realized immediately as a nonlinear availability or financial cost. Each decision determines the transitional probabilities for the next aircraft state and each action prescribes a policy for stochastic decision processes which impose constraints on the model. The optimal policy is realized by minimizing the long run expected average cost. The architecture may be any mathematical and/or logical instantiation and does not require neural networks – although neural networks are wholly realizable within this framework. A benefit of this approach is a fully transparent and explainable model unlike the "black boxes" of mainstream AI.

The paper also explains how this model may be extended to be a more effective decision-making/prescriptive analytics tool for aviation maintainers and fleet management. Further extensions include capital equipment in general, whether military, industrial, or commercial.

ABOUT THE AUTHORS

Randal Allen is the Chief Scientist of Lone Star Analysis. He is responsible for applied research and technology development across a wide range of M&S disciplines and manages intellectual property. He maintains a CMSP with NTSA. He has published and presented technical papers and is co-author of the textbook, "Simulation of Dynamic Systems with MATLAB and Simulink." He holds a Ph.D. in Mechanical Engineering (University of Central Florida), an Engineer's Degree in Aeronautical and Astronautical Engineering (Stanford University), an M.S. in Applied Mathematics and a B.S. in Engineering Physics (University of Illinois, Urbana-Champaign). He serves as an Adjunct Professor/Faculty Advisor in the MAE department at UCF where he has taught over 20 aerospace-related courses.

Zachry Engel is a Research and Development Engineer for Lone Star Analytics. Responsible for the development of new technologies for Lone Star, Zachry's primary interests within the company involve studying Artificial Intelligence and Optimization algorithms. He is also a member of the Technologies Futures team which helps guide the future technological interests at Lone Star. Zachry completed his doctorate in Mathematics in August of 2020 at the University of Texas at Arlington. Zachry's research interests involved Biostatistics and Optimization algorithms. His studies include receiving a bachelor's degree in Mathematics in May of 2016 from the University of Texas at Arlington.

Eric Haney is the Chief Technology Officer for Lone Star Analytics and previously served as Deputy CTO and Lead Analyst at the company. In his current role, Eric is responsible for guiding the technical direction of the company. He manages the development and roll-out of Lone Star's core software applications, TruNavigator™, AnalyticsOS™, and TruPredict™. His principal goal is to connect individuals with the information required to make pragmatic, objective decisions. Eric completed his doctorate in Aerospace Engineering at the University of Texas at Arlington culminating in his dissertation, *Data Engineering in Aerospace Systems Design & Forecasting*. His academic research team's collaborations with DARPA and NASA have been published in The Aeronautical Journal.

Evolved AI for Model-Based Reinforcement Learning

Randal Allen
Lone Star Analysis
Orlando, FL
rallen@lone-star.com

Zachry Engel, Eric Haney
Lone Star Analysis
Addison, TX
zengel@lone-star.com, ehaney@lone-star.com

MOTIVATION

Artificial Intelligence (AI) that teaches itself to achieve a goal is the next big thing, but most companies don't know how or when to apply it (Hume, Taylor 2021). Unlike supervised and unsupervised learning which are typically static optimization applications, reinforcement learning operates in a dynamic environment. To apply reinforcement learning, (Mah 2021) suggests businesses start with a list including a sequence of frequent actions with the opportunity for feedback. Brig. Gen. Eric Austin (director USMC Capabilities Development Directorate) said other key areas of AI development include support for maintenance and improving logistics (Tadjdeh 2021). These are motivations for selecting aviation maintenance as an application of model-based reinforcement learning described within this paper.

Motivation for applied research of *reinforcement learning* alone emanates from an interview with NDIA's Senior Fellow for AI. When asked how AI has transformed, Shane Shaneman (strategic director of national security and defense at Carnegie Mellon University) said, "…you've seen continued evolutions of both the algorithms and the framework and also new styles of machine learning. Of course, going…into new areas of both unsupervised as well as *reinforcement learning*…" (Tadjdeh, 2020).

Motivation for *model-based* reinforcement learning stems from the community's overall desire for novel approaches to machine learning beyond neural nets and deep learning. (Marcus 2020) identifies requirements necessary for moving toward what he calls Robust AI – a hybrid, knowledge-driven, cognitive *model-based* approach. A hybrid approach incorporates symbolic algorithms into the model architecture. Algorithms like these are prevalent in everyday computer operating systems and applications we fully trust.

Our Evolved AI approach takes hybrid and physics-based architectures to another level by blending data and causality. Afterall, "If you have insight into a problem that can help to solve it, then by all means, use it!" (Heaton, 2013). We have previously researched and reported on machine learning (supervised regression and unsupervised clustering) following this novel approach. So it's naturally motivating to follow up with research of reinforcement learning.

BACKGROUND

There are three fundamental types of machine learning methods, supervised learning (Allen, 2019), unsupervised learning (Allen 2020), and reinforcement learning – selection and implementation of which depends on how data is being manipulated. In the case of *supervised learning*, a model is trained on known input and known output data to predict future outputs. There are two subsets to supervised learning: <u>regression</u> techniques for continuous prediction and <u>classification</u> techniques for discrete prediction. In the case of *unsupervised learning*, clustering techniques are used to identify patterns and/or structures in the input data. There are two subsets of unsupervised learning: <u>hard</u> clustering where each data point belongs to only one cluster and <u>soft</u> clustering where each data point can belong to more than one cluster. There are two types of reinforcement learning: <u>model-free</u> and <u>model-based</u>. With model-free reinforcement learning, the agent knows nothing about its environment and attempts to find the optimal outcome by performing an exhaustive search. With model-based approaches, the practitioner provides a framework for the agent to operate within thereby assisting the search for an optimal outcome. While model-free techniques are currently more popular, the focus of this paper is on *model-based* reinforcement learning driven by the need for a hybrid/multi-model approach (Hao 2020) as required for Robust AI (Marcus 2020).

EVOLVED AI

Figure 1 shows the architecture of Evolved AI which functions as follows: signal #2 is sent to the mathematical model yielding output signal #3. The error signal #4, which is the difference between the feedforward signal #1 and the output signal #3, is minimized. Minimization of error signal #4 is achieved through optimization techniques (Allen, 2019) by adjusting parameters within the mathematical model. By minimizing the error, signal #3 will match signal #1 and thus, characterize the mathematical model.

As a practical example, consider digital filter design where a Bode plot is given with a desired amplitude response. The task is to fit the desired response with a 5^{th}-order polynomial (mathematical model). By comparing the desired signal #1 with output signal #3 via an error function, e.g., a Euclidean norm, optimization (minimization of the norm) will force the coefficients of the polynomial to produce the desired response, known as a matched filter.

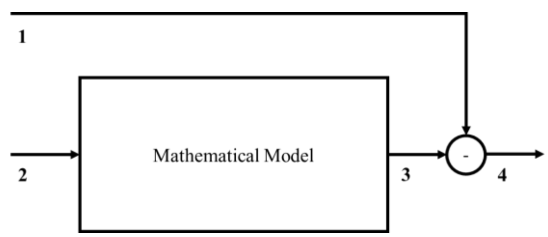

Figure 1 - Evolved AI's Architecture

The mathematical model may be generic or specific, depending on the application. If available, a practitioner with area expertise should incorporate *a priori* knowledge into the design of the mathematical model. It is important to point out that the mathematical model can include nonlinearities, nonconvexities, and mathematical and/or logical discontinuities. Optimization of nonlinearities is not unique, but nonconvexity certainly reduces the set of systems which provide solutions. When *mathematical* discontinuities are included, the number of systems providing solutions reduces further. Yet, even in this case, current machine learning techniques find it difficult to approximate this type of discontinuity. When considering *logical* discontinuities, e.g., Boolean logic gates like AND, OR, NOR, etc., the typical machine learning solution is to implement a perceptron for a single logic gate. However, Evolved AI allows the practitioner to utilize the actual logic gate(s) explicitly even extending to compound discontinuities such as truth tables, multi-input minimum/maximum, and if-then logic.

Since model-based reinforcement learning requires the practitioner to provide a framework, a mathematical (cost) model will be developed in the next section with application to aviation maintenance. The goal is to search for an optimal maintenance policy given various maintenance approaches depending on the state of the aircraft. Since no neural nets are used to solve this problem, the opaqueness of machine learning "black boxes" is removed so stakeholders can easily interpret the results and fully explain how the transparent system works.

AVIATION MAINTENANCE

Challenges in aviation maintenance include "just in time" maintenance and logistic support chain requirements. Optimizing the correct level of investment in repairable assets (spare parts) while minimizing storage footprint requirements at multiple sites is highly desirable. Particularly challenging is accounting for the number of spare assets needed when there is a lack of availability due to shortages in the industrial supply chain, an inaccurate forecast, and/or uncertain maintenance turn-around times. This multi-dimensional, time-varying problem is currently addressed with human maintenance schedulers and heuristics – this can lead to situations with ambiguous actions and unintended long-term consequences.

The solution to this problem is to apply model-based reinforcement learning. Aircraft states are defined according to Equipment Operational Capability (EOC) codes which classify the degradation of equipment mission capability. Based on the EOC codes, actions are selected from a set of possible policy choices. The consequence of an action is realized immediately as a nonlinear availability or financial cost. Each decision determines the transitional probabilities for the next aircraft state. Each action prescribes a policy for stochastic decision processes – which impose constraints on the model. The optimal policy is realized by minimizing the expected average cost in the long run.

Each aircraft platform and/or Weapons Replaceable Assembly (WRA) are assigned a OPNAV 5442 Mission Essential Subsystem Matrix (MESM). This matrix provides the EOC codes for a particular aircraft/platform. These codes range from "A" to "Z" with "A" being fully operational and "Z" being not operational at all. Normally around the "L-Y" codes the system or subsystem is severely impacted and could be restricted in use for example a "L" code could be assigned to an aircraft with an Anti-Icing problem and could be restricted from flying in "Known Icing Conditions." EOC codes may be linked back to supply priority/urgency codes which determine the priority placed upon an asset in need of repair or replacement.

Solution Formulation

This section explains the different aspects of the solution formulation, including the conditional states of the aircraft, which actions/policies may be adopted, and costs associated with each action.

The aircraft can be in one of four states: Brand New, Full Mission Capable (FMC), Partial Mission Capable (PMC) due to a single or multiple EOC codes, or Non-Mission Capable (NMC), as summarized in Table 1.

Table 1 –Aircraft States

State	Explanation
0	Brand New
1	Full Mission Capable (FMC)
2	Partial Mission Capable (PMC)
3	Non-Mission Capable (NMC)

At the risk of oversimplification for the sake of clarity, assume maintenance personnel can take only three different courses of action: (1) Do Nothing, (2) Overhaul, or (3) Replace the aircraft, summarized in Table 2. An overhaul might be required if there is a discrepancy with the aircraft, for instance corrosion cracks. Replacing the aircraft is required when the aircraft is beyond economic repair.

Table 2 –Maintenance Actions

Action	Explanation
1	Do Nothing
2	Overhaul
3	Replace

If the Do-Nothing policy is adopted, then the mission capability state of the aircraft will degrade over time. While there is a natural progression from Brand New through each state to NMC, there is some probability of skipping states depending on discrepancies. These probabilistic state transitions are shown in the state transition matrix of Table 3. All of the probabilities in the following tables have been reasonably assumed. For more precise probabilistic entries, one could access historical aircraft maintenance records.

Table 3 – "Do Nothing" Policy

Do Nothing State Transition Matrix				
State	0	1	2	3
0	0	7/8	1/16	1/16
1	0	3/4	1/8	1/8
2	0	0	1/2	1/2
3	0	0	0	1

As shown in Table 3, if the aircraft is initially in state 0 (Brand New), there is an 87.5% probability of transitioning to state 1 (FMC), a 6.25% probability of transitioning to state 2 (PMC), and a 6.25% probability of transitioning to state 3 (NMC). It is assumed the Brand-New aircraft transitions to the FMC instantaneously – analogous to driving a new automobile off the lot. If the aircraft is initially in state 1 (FMC), there is a 75% probability of remaining in state 1 (FMC), and a 12.5% probability of transitioning to state 2 (PMC) or state 3 (NMC). There is 0% probability of transitioning to state 0 (Brand New) because of the Do-Nothing policy. If the aircraft is initially in state 2 (PMC), there is a 50% probability of remaining in state 2 (PMC), and a 50% probability of transitioning to state 3 (NMC). Due to the Do-Nothing policy, there is 0% probability of transitioning back to prior states. Finally, if the aircraft is NMC, it remains in that state due to the Do-Nothing policy.

Now, assume the "Overhaul" policy takes the aircraft to state 1 (FMC) regardless of its prior state. These 100% probabilities are shown in Table 4. Note that this policy allows for overhauling a Brand-New aircraft and placing into the FMC state – which is absurd. Not to worry, the cost of doing so will prohibit the solution from selecting this as an option. Additionally, it doesn't make sense to overhaul an FMC aircraft only to have it be in the same FMC state. Again, the cost of this effort (to be described shortly) will prohibit it as an option.

Table 4 – "Overhaul" Policy

Overhaul State Transition Matrix				
State	0	1	2	3
0	0	1	0	0
1	0	1	0	0
2	0	1	0	0
3	0	1	0	0

Finally, consider the "Replace" policy, when the aircraft is sent to state 0 (Brand New) independent of its prior state. These 100% state transitions are captured in Table 5. Similar to the "Overhaul" policy, the "Replace" policy has three situations that don't make sense, especially replacing a Brand-New aircraft with a Brand New one. Again, costs will render these options out of the running.

Table 5 – "Replace" Policy

Replace State Transition Matrix				
State	0	1	2	3
0	1	0	0	0
1	1	0	0	0
2	1	0	0	0
3	1	0	0	0

Next, equations relating the conditional probabilities for a state (s), given an action (a), P(s|a), are formulated.

Begin with the Brand-New state 0 and the probabilities for each action: Do Nothing 1, Overhaul 2, and Replace 3.

$$P_{01} + P_{02} + P_{03} = 0P_{01} + 0P_{02} + 1P_{03} + 0P_{11} + 0P_{12} + 1P_{13} + 0P_{21} + 0P_{22} + 1P_{23} + 0P_{31} + 0P_{32} + 1P_{33}$$

The left side of this equation represents the probabilities of being in the Brand-New state and taking any of the three actions. The right side represents the probabilities of each policy returning any given state (0 through 3) to the Brand-New state (0). Notice the column of P_{01}, P_{11}, P_{21}, P_{31} corresponds to the 1st column of the *Do Nothing* policy; the column of P_{02}, P_{12}, P_{22}, P_{32} corresponds to the 1st column of the *Overhaul* policy; and the column of P_{03}, P_{13}, P_{23}, P_{33} corresponds to the 1st column of the *Replace* policy. The zero coefficients are intentionally retained for clarity.

Likewise, the equation representing FMC (state 1) and the probabilities for each action (the 2nd column of each corresponding policy) result in the following equation:

$$P_{11} + P_{12} + P_{13} = \frac{7}{8}P_{01} + 1P_{02} + 0P_{03} + \frac{3}{4}P_{11} + 1P_{12} + 0P_{13} + 0P_{21} + 1P_{22} + 0P_{23} + 0P_{31} + 1P_{32} + 0P_{33}$$

The equation representing PMC (state 2) and the probabilities for each action (the 3rd column of each corresponding policy) result in the following equation:

$$P_{21} + P_{22} + P_{23} = \frac{1}{16}P_{01} + 0P_{02} + 0P_{03} + \frac{1}{8}P_{11} + 0P_{12} + 0P_{13} + \frac{1}{2}P_{21} + 0P_{22} + 0P_{23} + 0P_{31} + 0P_{32} + 0P_{33}$$

Finally, the equation representing NMC (state 3) and the probabilities for each action (the 4th column of each corresponding policy) result in the following equation:

$$P_{31} + P_{32} + P_{33} = \frac{1}{16}P_{01} + 0P_{02} + 0P_{03} + \frac{1}{8}P_{11} + 0P_{12} + 0P_{13} + \frac{1}{2}P_{21} + 0P_{22} + 0P_{23} + 0P_{31} + 0P_{32} + 0P_{33}$$

Simplifying the set of equations in the following format prepares them as equality constraints for solving the model-based reinforcement learning problem.

$$P_{01} + P_{02} - P_{13} - P_{23} - P_{33} = 0$$
$$\frac{1}{4}P_{11} + P_{13} - \frac{7}{8}P_{01} - P_{02} - P_{22} - P_{32} = 0$$
$$\frac{1}{2}P_{21} + P_{22} + P_{23} - \frac{1}{16}P_{01} - \frac{1}{8}P_{11} = 0$$
$$P_{31} + P_{32} + P_{33} - \frac{1}{16}P_{01} - \frac{1}{8}P_{11} - \frac{1}{2}P_{21} = 0$$

Finally, costs need to be considered. Again, reasonable assumptions will be made. For more precise costs, consult aircraft maintenance records. If the aircraft is in the Brand-New state (0), no maintenance is required (Do Nothing, action 1) hence there is no cost. As mentioned previously, it would be absurd to overhaul or replace a Brand-New aircraft, so these probabilities will be omitted, i.e., $P_{02} = P_{03} = 0$. Now if the aircraft is in the FMC state (1) and no maintenance is performed (Do Nothing, action 1), the aircraft will become defective over time perhaps at a cost of \$1M. If the aircraft is in the PMC state (2) and no maintenance is performed (Do Nothing, action 1), the aircraft will become more defective perhaps increasing costs to \$3M. As a simplification, assume the aircraft is only overhauled (action 2) if it is PMC (state 2) at a cost of \$4M and hence, $P_{12} = P_{32} = 0$. Finally, let the cost to replace (action 3) the aircraft be \$6M, which is the case regardless of the state of the aircraft. The astute reader will notice the final case is when the aircraft is NMC and the Do-Nothing policy is adopted. It doesn't cost anything to Do Nothing with an NMC aircraft. Since there is no contribution to the overall cost, $P_{31} = 0$. The costs associated with each action, corresponding to a particular state, are summarized in Table 6.

Table 6 – Cost Matrix

State	Action	Cost ($M)
0	1	0
1	1	1
2	1	3
2	2	4
1	3	6
2	3	6
3	3	6

From Table 6, the total cost ($M) is the summation of the individual costs

$$0P_{01} + 1P_{11} + 3P_{21} + 4P_{22} + 6P_{13} + 6P_{23} + 6P_{33}$$

The model-based reinforcement learning problem may now be cast. The objective is to minimize the total cost subject to the equality constraints, with the inclusion of an additional equality constraint accounting for the sum of the probabilities to be unity. Recall with model-based approaches, the practitioner provides a framework for the agent to operate within thereby assisting the search for an optimal outcome. Here, the equality constraints provide that framework, assisting in the search for an optimal maintenance policy. In mathematical parlance,

$$\{min\}\ 1{,}000{,}000P_{11} + 3{,}000{,}000P_{21} + 4{,}000{,}000P_{22} + 6{,}000{,}000(P_{13} + P_{23} + P_{33})$$

Subject to:

$$P_{01} - P_{13} - P_{23} - P_{33} = 0$$
$$\tfrac{1}{4}P_{11} + P_{13} - \tfrac{7}{8}P_{01} - P_{22} = 0$$
$$\tfrac{1}{2}P_{21} + P_{22} + P_{23} - \tfrac{1}{16}P_{01} - \tfrac{1}{8}P_{11} = 0$$
$$P_{33} - \tfrac{1}{16}P_{01} - \tfrac{1}{8}P_{11} - \tfrac{1}{2}P_{21} = 0$$
$$P_{01} + P_{11} + P_{21} + P_{22} + P_{13} + P_{23} + P_{33} - 1 = 0$$

The Evolved AI model is shown in Figure 2. The probabilities are shown on the left, individual costs are shown in the ovals, and total cost is the summation shown on the right.

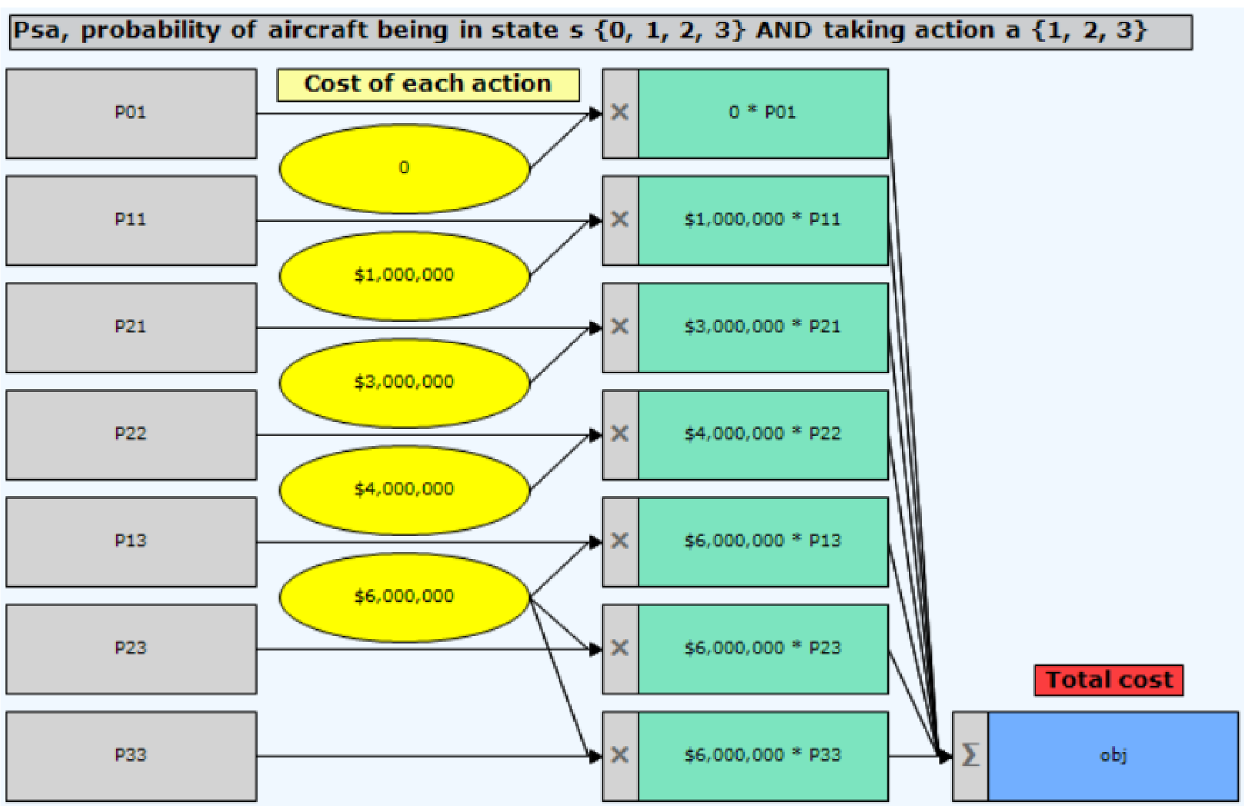

Figure 2 – Model-Based Reinforcement Learning in Evolved AI

Upon execution of the model, the optimal policy probabilities are summarized in Table 7. First, it is seen that the probabilities sum to unity.

Table 7 – Probabilities, States, and Optimal Policies

Probability	Value	State	Policy
P_{01}	2/21	0, Brand New	1, Do Nothing
P_{11}	5/7	1, FMC	1, Do Nothing
P_{13}	0		
P_{21}	0		
P_{22}	2/21	2, PMC	2, Overhaul
P_{23}	0		
P_{33}	2/21	3, NMC	3, Replace

Second, interpretation of the non-zero probabilities provide the optimal policy. Since P_{01} is non-zero, instructions are for state 0, perform action 1, i.e., for a Brand-New aircraft, Do Nothing. Likewise, for an FMC aircraft, Do Nothing; for a PMC aircraft, Overhaul; for an NMC aircraft, Replace. Hance, the optimal policy is provided for each state pf the aircraft. The (minimum) total cost for adopting this policy is $1,666,667.

SUMMARY

Aircraft states are defined according to Equipment Operational Capability (EOC) codes which classify the degradation of equipment mission capability. Based on the EOC codes, actions are selected from a set of possible choices. The consequence of an action is realized immediately as a nonlinear availability or financial cost. Each decision determines the transitional probabilities for the next aircraft state. Each action prescribes a policy for stochastic decision process – which impose constraints on the model. The optimal policy is realized by minimizing the expected average cost in the long run.

This approach to mathematically assessing the maintenance policy of an aircraft, given its contextual state, should be compared against the status quo for aviation maintenance. Currently, aircraft are typically under a preventative maintenance schedule where defined tasks are performed on rotating intervals of time and/or utilization. The consideration of costs, current status of the aircraft, or the dynamic operating environment of business are not taken into account, or are left to human subjectivity. This leaves actions open to interpretation and omits major business impacts. In addition, decisions are made on an asset-by-asset basis and are not related to their impacts on overall fleet availability. By imparting objective analysis through model-based reinforcement learning, these multi-faceted decisions can be made more consistent and reflective of overall operations.

FUTURE WORK

This high-level model may also be extended to be a more effective decision-making/prescriptive analytics tool for aviation maintainers and fleet management. The courses of action required to satisfy the Summary statement in the preceding section will vary greatly and will be driven by large uncertainty related to the following conditions at the minimum 1) next intended mission/use of the aircraft, 2) availability of materiel resources, and 3) availability of down time to complete the repair and the skillset of the artisan required. Variability substantially increases when the EOC code and mission-driven decisions directly impact Survivability or Lethality of the mission aircraft such as Electronic Countermeasure or Deceptive Electronic Countermeasure (ECM/DECM) systems designed to increase survivability in high threat situations. These conditions can be included into the model since Evolved AI naturally incorporates uncertainty and the cost function can include a multitude of uncertain considerations. Furthermore, aircraft states could be expanded by individual EOC code for higher fidelity and probabilistic cost metrics could be obtained to address airframe-centric maintenance requirements. Additional extensions include capital equipment in general, whether military, industrial, or commercial.

ACKNOWLEDGEMENTS

The authors thank Rob Harrill (Vice President, Operational Optimization Solutions) for providing the context and details of naval aircraft maintenance thus providing context for a practical example suitable for an I/ITSEC audience.

REFERENCES

1. Allen, R. (2019), Adaptive Nonconvex Optimization for AI and Machine Learning, I/ITSEC paper 19109.
2. Allen, R. (2019), Evolved AI - Interpretable Network Architectures for Machine Learning, I/ITSEC paper 19149.
3. Allen, R. (2020), Evolved Artificial Intelligence for Stochastic Clustering Unsupervised Learning, I/ITSEC paper 20258.
4. Hao, K., (2020), A Debate Between AI Experts Shows a Battle Over the Technology's Future, MIT Technology Review, https://www.technologyreview.com/s/615416/ai-debate-gary-marcus-danny-lange.
5. Heaton, J. (2013), Artificial Intelligence for Humans Volume 1: Fundamental Algorithms, p.155
6. Hume K. and Taylor, M. (2021), Why AI That Teaches Itself to Achieve a Goal Is the Next Big Thing, Harvard Business Review, https://hbr.org/2021/04/why-ai-that-teaches-itself-to-achieve-a-goal-is-the-next-big-thing
7. Mah, P. (2021), Why Reinforcement Learning Matters in Business, AI/ML CDO Trends, https://www.cdotrends.com/story/15555/why-reinforcement-learning-matters-business
8. Marcus, G. and Davis, E., (2019), Rebooting AI: Building Artificial Intelligence We Can Trust, Pantheon.
9. Marcus, G., (2020), The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence, https://arxiv.org/pdf/2002.06177.
10. Naval Supply Systems Command, (1998), MILSTRIP/MILSTRAP Desk Guide, NAVSUP Publication 409, Revised 06/98, COG | Stock No. 0530-LP-011-2240.
11. Commander, Naval Air Force, Pacific, (2014), MH-53E Mission Essential Subsystem Matrix (MESM).
12. Commander, Naval Air Force, U.S. Pacific Fleet, (2019), F/A-18A/B/C/D Mission Essential Subsystem Matrix (MESM).
13. Commander, Naval Air Force, Atlantic, (2020), CH-53E Mission Essential Subsystem Matrix (MESM).
14. Sutton, R.S., and Barto, A.G., (2018), Reinforcement Learning: An Introduction, MIT Press.
15. Tadjdeh, Y., Interview with NDIA's Senior Fellow for AI, National Defense Magazine, 10 Jan 2020 https://www.nationaldefensemagazine.org/articles/2020/1/10/interview-with-ndias-senior-fellow-for-ai.
16. Tadjdeh, Y., Marines Lack Trust in Artificial Intelligence, National Defense Magazine, April 2021.