Energy Management of a Hybrid Electric Vehicle using Model-Based Reinforcement Learning

Project Report

Prepared in partial fulfilment of the

Study Project (ME F266)

Prepared By

Jainil Dharmil Shah – 2020A4PS0504P

Submitted To

Dr Saket Verma,

Department of Mechanical Engineering,

Birla Institute of Technology and Science, Pilani

Acknowledgements

I would like to acknowledge my deep sense of gratitude to Dr Saket Verma for providing me this opportunity to work on this project of ‘Energy Management Strategy of Fuel Cell Electric Vehicles Using Model-Based Reinforcement Learning’. The role of a mentor played by him during the entire duration of this project played a pivotal role in developing new skills, gaining lots and lots of knowledge and keeping me motivated.

I would next like to sincerely thank Dr Kuldeep Kumar, the most valuable driving force behind this project, without whom, this project could’t have been taken forward. His valuable inputs during every project meet guided me on what had to be done next and how the project had to be approached in a step by step manner.

My immense gratitude is also due, to Prof. M.S. Dasgupta, HoD, Department of Mechanical Engineering, BITS Pilani – Pilani Campus and also Prof. Suresh Gupta, Associate Dean, Academic Undergraduate Studies Division, BITS Pilani – Pilani Campus for providing UG students to undertake such project type courses as a part of their academic curriculum, thus giving research exposure to students at an early stage.

Table of Contents

Abstract

Fuel cell electric vehicles employ fuel cells as their primary power source, with an electric motor driving the vehicle and a secondary power source that stores regenerative braking energy and supports driving. An energy management system for the optimal distribution of power among these fuel cells and electric batteries is required to reduce hydrogen fuel consumption by effectively utilising these fuel cells and electric batteries. In this study, an attempt was made to use Model Based Reinforcement Learning for energy management. Reinforcement learning is used as a learning algorithm for optimal control of a fuel-cell electric vehicle; the model is initially unknown, but it is taught with data from experiences as the learning process advances. The control policy is then optimised using reinforcement learning for the environment of the driving cycle profile.

Emergence of Hybrid Electric Vehicles

Automobiles have contributed significantly to the development of modern society by addressing the desire for greater mobility in daily life. The advancement of ICE has had a significant impact on the vehicle industry. However, enormous volumes of hazardous emissions such as carbon dioxide (CO_2), carbon monoxide (CO), nitrogen oxides (NO_x), unburned hydrocarbons (HCs), and other pollutants have resulted in pollution, global warming, and ozone layer depletion. These emissions endanger both the environment and human life. Furthermore, because petroleum resources are limited, petroleum usage must be decreased. Alternative transportation technology, which uses ICE as the major power source and batteries/electric motor as the peaking power source, is one noteworthy answer to these problems. This notion has given rise to new modes of transportation such as electric vehicles (EVs), battery electric vehicles (BEVs), hybrid electric vehicles (HEVs), and plug-in hybrid electric vehicles (PHEVs), all of which are clean, cost-effective, efficient, and environmentally friendly.

High-efficiency electric motors and controllers enable the EVs, which are powered by alternate energy sources. EVs are a clean, efficient, and environmentally beneficial mode of urban transportation, but their range is restricted. HEVs were created as a result of the increased battery costs, limited driving range, and performance of EVs. During vehicle propulsion, HEVs use both an electric motor and an internal combustion engine (ICE). It combines the benefits of both ICE and EV vehicles while eliminating their drawbacks. In HEVs, the battery serves as a backup power system for the ICE during vehicle propulsion, reducing liquid fuel usage and hazardous emissions. Ferdinand Porsche created the first gasoline-electric hybrid car, the Lohner Porsche Mixte Hybrid, in 1901.

However, HEV still relies on fossil fuel, which means the Green House Gases (GHG) and environmental pollutants will be discharged inherently. As a result, a Fuel Cell Electric Vehicle (FCEV) without an engine is suggested. The energy management problem, aimed at reducing environmental impact, is one of the most important considerations among all technologies involved in FCHEV.

Architecture of Hybrid Electric Vehicles

HEVs are mainly classified into three main categories: 1.) Series Hybrid 2.) Parallel Hybrid 3.) Series-Parallel (Power-split) Hybrid.

The series configuration consists of an electric motor with an ICE without any mechanical connection between them. When the battery is insufficient to move the car, ICE is used to power an electric generator rather than driving the wheels directly. Series hybrids have only one driving train, but all operations require two different energy conversion systems. Gasoline to electricity and electricity to drive wheels are the two energy conversion procedures. Real-time monitoring and improvement of fuel economy and power source performance.

In a parallel setup, a single electric motor and an internal combustion engine are fitted in such a way that they can operate the vehicle separately or jointly. Parallel hybrids allow both power sources to operate at the same time for maximum efficiency. While this method improves economy and performance, it also increases the complexity and cost of the transmission and drive train. The parallel configuration is more difficult than the series configuration, but it has advantages.

Power split hybrid has a combination of both series and parallel configuration in a single frame. The engine and battery can power the vehicle separately or jointly in this form, and the battery can be charged at the same time through the engine. It basically extends a hybrid vehicle's all-electric range (AER). The power-split design, which can be divided into two modes, is currently the most popular architecture. The two modes are as follows: 1.) one (single) mode and 2.) two (dual) modes. Single mode contains one planetary gear set (PGS) and dual mode contains two PGS which are required for a compound power split. It is further classified into three types: 1.) Input Slip 2.) Output Slip 3.) Compound Split. These splits by the method of output of delivery.

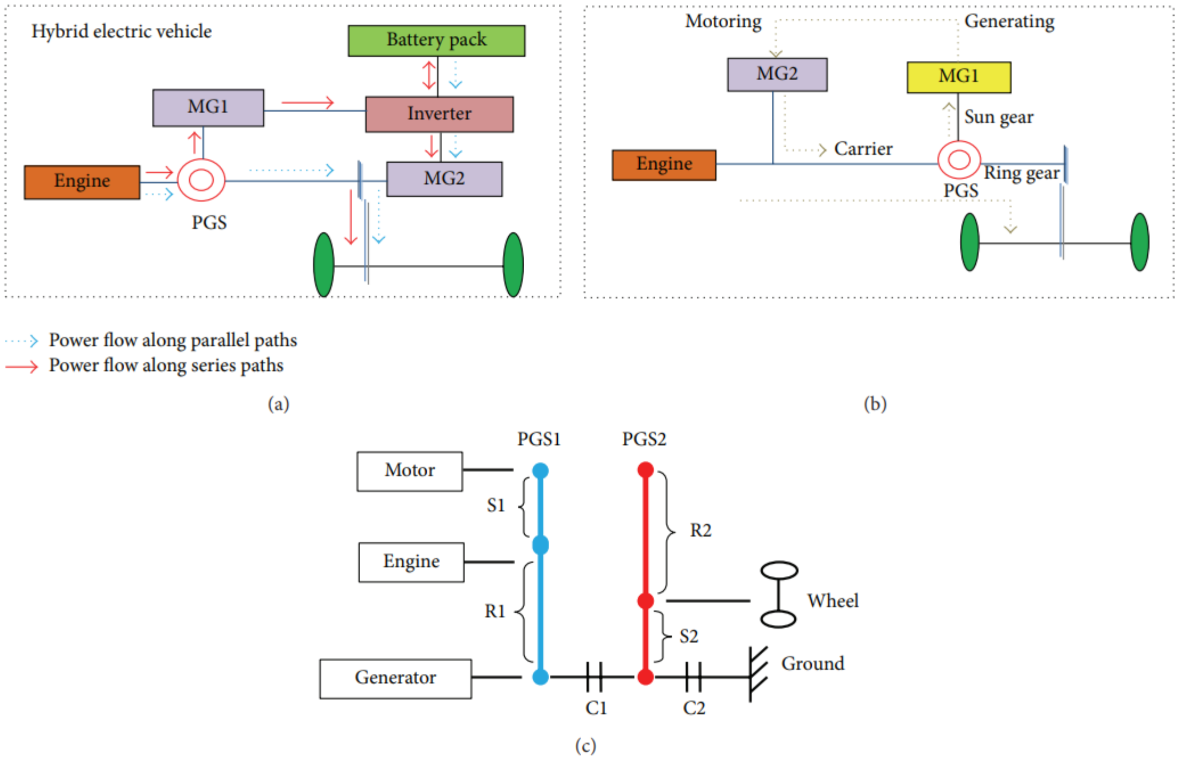


Figure 1: Power-split configurations: (a) input split, (b) output split, and (c) compound split.

Vehicle Dynamics

The vehicle's speed is determined by the traction, resistive, and braking forces at work. The model is based on the longitudinal balance of forces. The DC bus PL's power demand is calculated by adding the power train losses from the wheels to the electric motor's input (P). The losses are divided into two categories: transmission system losses and motor/inverter losses. The driving cycle is typically a matrix with time and speed indices (v). The total forces (F_T) operating on the wheel during each time step must first be calculated, and then the power demand at the wheel levels must be calculated (Pw). The aerodynamic force Fw, the force of rolling resistance or friction losses F_r, the force owing to tilt or load slope Fi, and the acceleration force Fa make up the entire sum of forces F_T acting on the wheels.

$F_{load} = 0.5\rho_a A_f C_d v^2 + C_r mvg\cos\theta + mvg\sin\theta$ **(1)**

ρ_a is the air density, A_f is the vehicle front area, Cd is the vehicle drag coefficient, Cr is the vehicle rolling resistance coefficient, m is the mass of the vehicle and θ is the slope.

Aerodynamic Force: $0.5\rho_a A_f C_d v^2$; Rolling Resistance: $Fr = C_r mg\cos\theta$;

Inclination Force: $Fi = mg\sin\theta v$; Acceleration Forces: $Fa = ma$

Sum of Forces: $F_T = Fr + Fi + Fa$ **(2)**

Power Demand at Wheels: $P_w = vF_T$ **(3)**

Power Demand at Motor : $P_w + P_{losses}$ **(4)**

The Electric Power Demand can be expressed as follows: $P_{elec} = \eta_{elec}^{-sgn(Tm)} \cdot Tm \cdot \omega_m$ **(5)** , where T_m is the motor torque, η_{elec} is the efficiency of the motor and the converter and ω_m is the motor speed.

Battery State-of-Charge (SOC) is expressed as follows:

$$SOC = -\frac{1}{2}\frac{Vocv - \sqrt{(Vocv)^2 - 4PbatRbat}}{QbatRbat} \quad \textbf{(6)}$$

where Vocv is the open-circuit voltage, Pbat is the battery power, and Rbat is the internal resistance of the battery, both of which can be expressed as a function of the battery SOC.

Introduction (Problem Overview)

In Hybrid Electric Vehicles (HEVs), the presence of an alternate energy source in addition to the Internal Combustion Engine (ICE) encourages an optimal power split between both for minimum fuel consumption and maximum power usage. As a result, compared to ICE-based cars/conventional automobiles, HEVs give higher fuel economy. Due to the availability of two power sources, an energy management technique is required to share power between them. The technique should be able to reduce fuel use while increasing power usage.

Energy management strategies are algorithms that determine how much power goes to the engine and how much goes to the motor in order to increase fuel economy and HEV performance. The fuel efficiency of hybrid electric vehicles is heavily influenced by energy management systems. They are critical in dividing power between two sources, mainly the engine and the battery. The intelligent distribution of power between these two improves fuel economy and regulates power flow. The amount of power divided between the engine and the motor is determined by the battery's state of charge (SOC), the amount of power required at the wheels, and the engine's operational range.

One of the most critical aspects defining a vehicle's fuel economy performance is its energy management strategy for hybrid electric vehicles (HEVs). The energy management strategy is a supervisory control approach for operating each power source by determining when and how much energy to utilise according to the driving environment. It coordinates several power sources, mainly fossil fuel energy and electric energy in HEVs.

Recent research has been carried out in order to forecast and use future driving circumstances. However, precisely predicting future driving speed profiles is difficult, and altering and learning the best control strategy based on the vehicle's changing driving conditions necessitates a complex algorithm and processing overhead. As a result of these issues, subsequent approaches have attempted to leverage machine learning to solve hybrid control challenges.

Reinforcement Learning (RL), a subject of machine learning that has received a lot of attention in recent years, has a framework that may be used to effectively regulate problems. RL is a sort of machine learning that was created using dynamic programming as a foundation. As a result, problems that were previously solved using DP, such as the HEV optimal control problem, can be applied to the control problem framework using RL. In fact, based on past studies on stochastic dynamic programming, these RL techniques have been applied to HEV control.

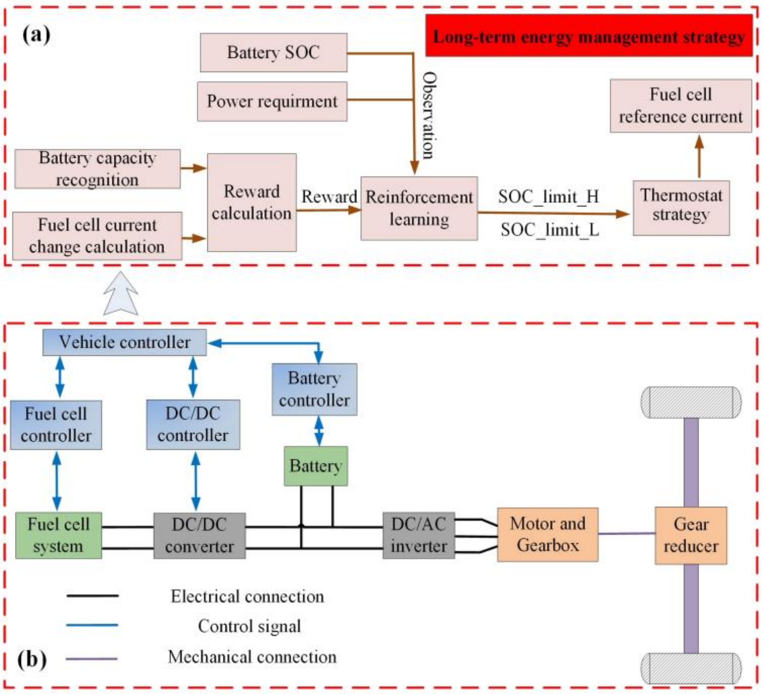


Figure 2: (a) Long-term energy management strategy; (b) powertrain structure of fuel cell/battery hybrid city bus

Classification of Energy Management Strategies

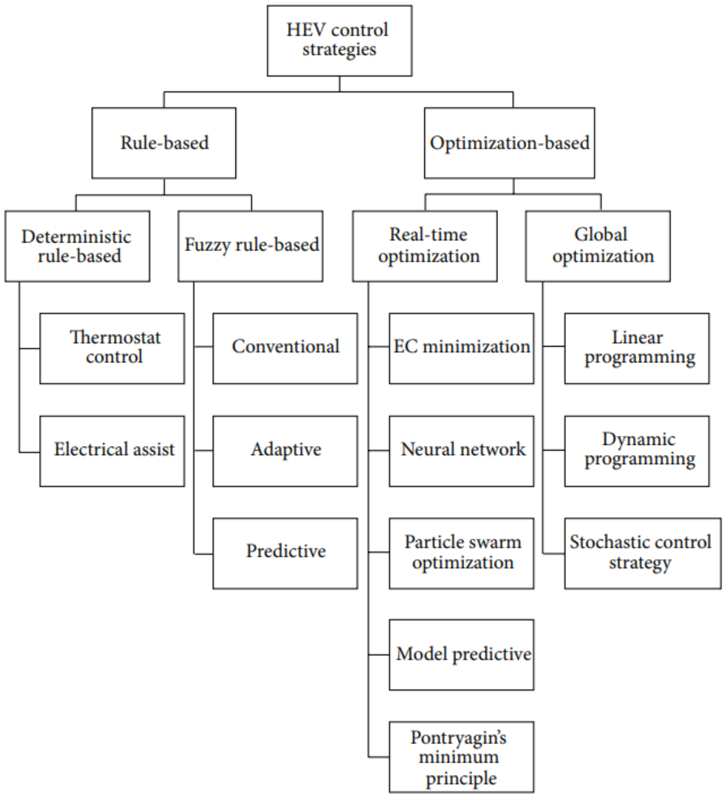


Figure 3: Classification of Control Strategies

Control strategies are broadly classified into rule-based and optimization-based control strategy and all other subcategories are classified based on these two main categories:

A. Rule based control strategy: Deterministic and fuzzy rule based:

A rule-based control strategy can be easily implemented with real-time supervisory control to manage power flow in a hybrid drive train. These are static controllers, with decisions based solely on instantaneous inputs. To meet the requirements of the driver and other components (electrical loads and batteries), the operating point of the components (ICE, traction motor, and generator) is calculated using rule tables or flowcharts. The creation of rule-based behaviour is made possible by fuzzy logic (FL). An expert's knowledge can be codified as a rule-base and used in decision-making. The key advantage of FL is that it can be modified and adapted to suit the situation, giving you more control.

B. Optimization based control strategy:

Energy Consumption Minimization Strategy (ECMS), Neural Network (NN), Particle Swarm Optimization (PSO), Model Predictive Control (MPC), and Pontryagin's Minimum Principle (PMP) have been employed in literature for power optimization in HEVs.

The overall fuel consumption of an electric motor is calculated using the ECMS, which is the sum of ICE and electric motor fuel consumption. This method calculates real-time equivalent fuel usage as a function of the current system measured parameters. Because of its adaptable structure, NN can be used in any control application. A well-designed network can adapt to any lookup table and update the table data through training.

PSO is a meta-heuristic method for searching a large number of candidate solutions. It can be utilised on irregular, noisy, temporal variant type optimization issues and does not require the optimization problem to be differentiable, as traditional optimization methods do. The upgraded PSO's multilevel hierarchical control approach can accurately determine the direction and quantity of energy flow in HEVs/PHEVs (Plug-in HEV).

The MPC approach is used to create a dynamic model of a process using system identification. It allows you to optimise your present time slot while also considering future time slots. MPC is capable of foreseeing future events and taking control actions in response. Because the number of nonlinear second order differential equations rises linearly with dimension in PMP, it requires less computational time to find an ideal route, but it may be a local rather than a global solution. PMP's ideal trajectory should be considered a global optimal trajectory under specific conditions.

C. Global Optimization based control strategies:

Linear Programming (LP), Dynamic Programming (DP), Stochastic Dynamic Programming (SDP), and GA are used to find the global solution of the optimization problems and are being widely used in hybrid vehicle applications.

The LP approach is used to solve the fuel consumption minimization problem, which is represented as a convex nonlinear optimization problem. It could lead to a global best solution. DP is a strategy for both mathematics optimization and computer programming. It's difficult to apply the "curse of dimensionality" to complex systems. Because the duty cycle must be known in advance, the DP method cannot be applied in real time. SDP refers to an optimization problem in which either the state or the choice is known in terms of a probability function. The stochastic optimum control issue necessitates the use of high-performance computing techniques.

GA stands for heuristic search algorithm, and it is used to find solutions to optimization and search problems. This is a branch of artificial intelligence inspired by Darwin’s Theory of Evolution. GA is a robust and viable approach that searches a large amount of space and quickly optimises the parameters using simple processes. As nonlinear, multi-modal, non-convex objective functions, they have been shown to be effective in solving complex engineering optimization issues. Without becoming caught in local optima, GA efficiently points to the global optima. The GA approach does not necessitate any strong assumptions or additional objective parameter information. GA has a lot of power when it comes to exploring the solution space.

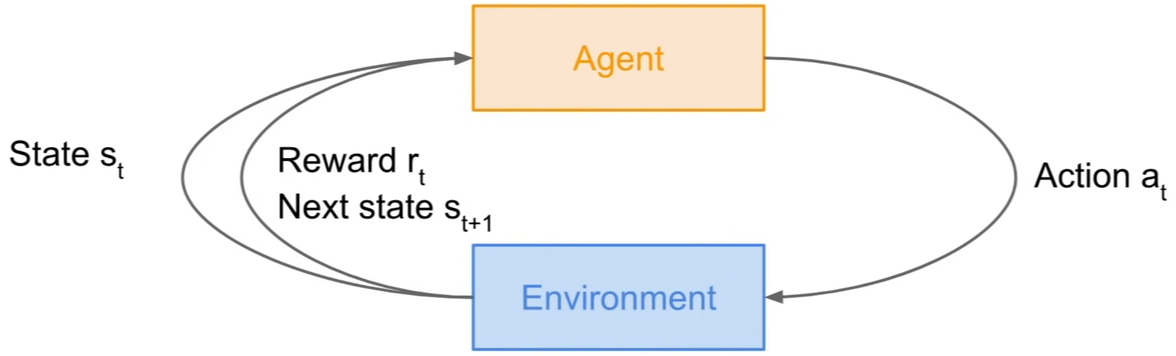
Reinforcement Learning

Mathematical Formulation of Reinforcement Learning:

In the reinforcement learning setup, we have an agent and we have an environment. The environment gives the agent a state. In turn the agent is going to take an action, and then the environment is going to give back a reward, as well as the next state. And so, this is going to keep going on in this loop, until the environment gives back a terminal state which then ends the episode.

Defined by: (S, A, R, P, γ)

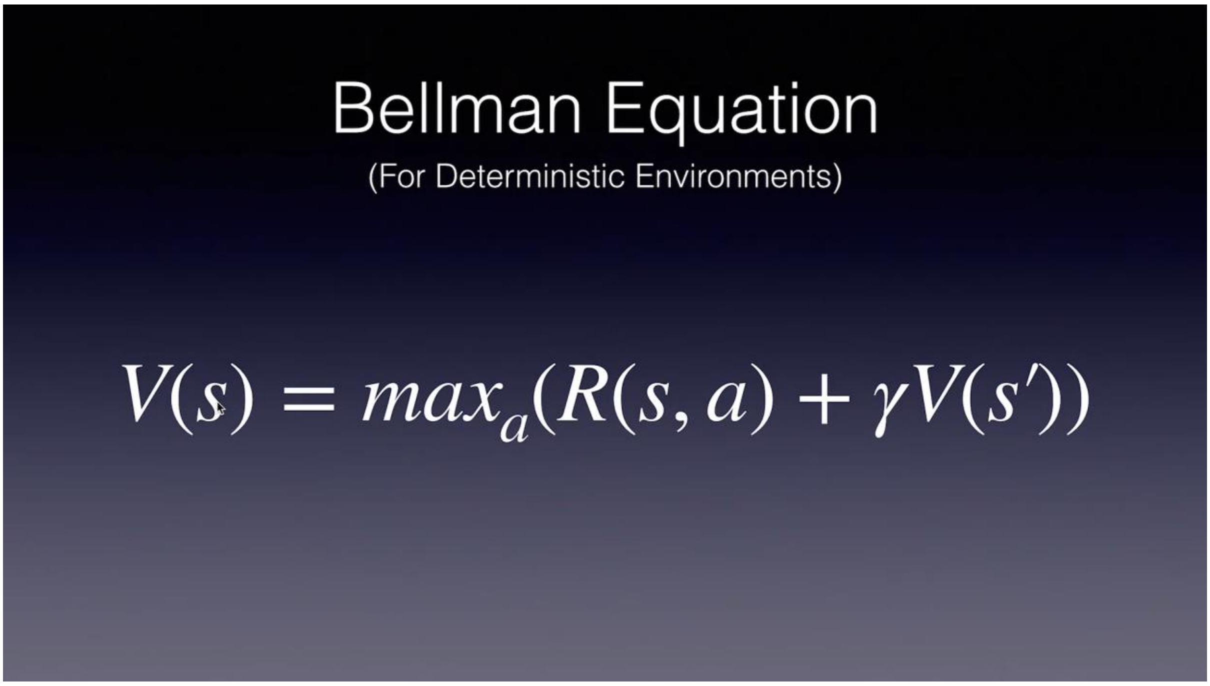
1. S -> Set of possible states
2. A -> Set of possible actions
3. R -> Distribution of reward given (state, action) pair
4. P -> Transition Probability, i.e. distribution over next state given (state, action) pair
5. γ -> Discount Factor



Dynamic Programming:

Richard Bellman used the term "dynamic programming" in 1940 to describe the process of solving problems that require successively finding the optimum decisions. DP is a computer programming method as well as a mathematical optimization method. It refers to breaking down a complex problem into smaller subproblems in a recursive manner in both contexts.

The essential foundation of this method is the notion of optimality. There are two approaches to approaching the optimal solution to the problem when you have a dynamical process and the accompanying performance function. The maximal principle of Pontryagin is one, and Bellman's dynamic programming is the other. It has the advantage of being applicable to both constrained and unconstrained as well as linear and nonlinear problems.



For the given project, Dynamic Programming (DP) is a well-known strategy for finding a global optimal solution. DP is based on the Bellman equation, which represents the optimization problem in a recursive form, and can solve the problem efficiently. However, the DP calculation is quite time consuming, making real-life application difficult. As the number of states increases, the complexity of calculation also increases rapidly owing to the so-called ‘‘curse of dimensionality’’.

Model-Based Reinforcement Learning:

Model-based reinforcement learning (MBRL) was examined for FCEVs in this study. MBRL uses an internal model to approximate the environment and the control behaviour can be learned through this model.

MBRL are more efficient than model-free approaches due to the following reason:

- In model-based learning, the optimal policy can be extracted from the calculation using the learned model.
- In the model-free method, trial-and-error is required to determine the optimal control policy, which is a time-consuming process.

Use of Reinforcement Learning Toolbox

Reinforcement Learning Toolbox provides an app, functions, and a Simulink block for training policies using reinforcement learning algorithms, including DQN, PPO, SAC, and DDPG. For sophisticated applications such as resource allocation, robotics, and autonomous systems, these principles can be used to create controllers and decision-making algorithms. The toolbox allows you to build policies and value functions with deep neural networks or look-up tables, and then train them in MATLAB or Simulink environments. One can test the toolbox's single- and multi-agent reinforcement learning algorithms or create your own. One may play with with hyperparameter settings, track training progress, and simulate trained agents interactively or programmatically using the app.

1.) <u>Train DQN Agent to Balance Cart-Pole System using Reinforcement Learning Toolbox in Matlab:</u>

A pole attached to an unactuated joint on a cart that runs over a frictionless track serves as the reinforcement learning environment in this case. The purpose of the training is to get the pole to stand up straight without falling over.

1.) Create Environment Interface

2.) Create DQN Agent

3.) Train Agent

4.) Simulate DQN Agent

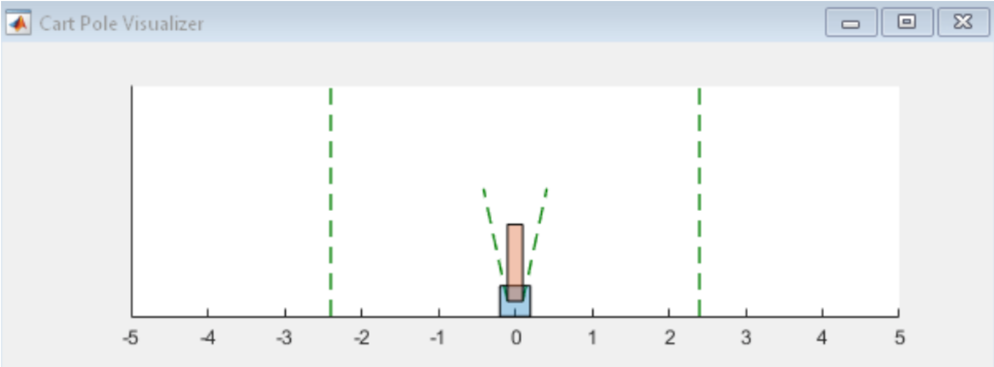


Figure 4: Visualization of Balancing of Cart-Pole System using Reinforcement Learning Toolbox in Matlab

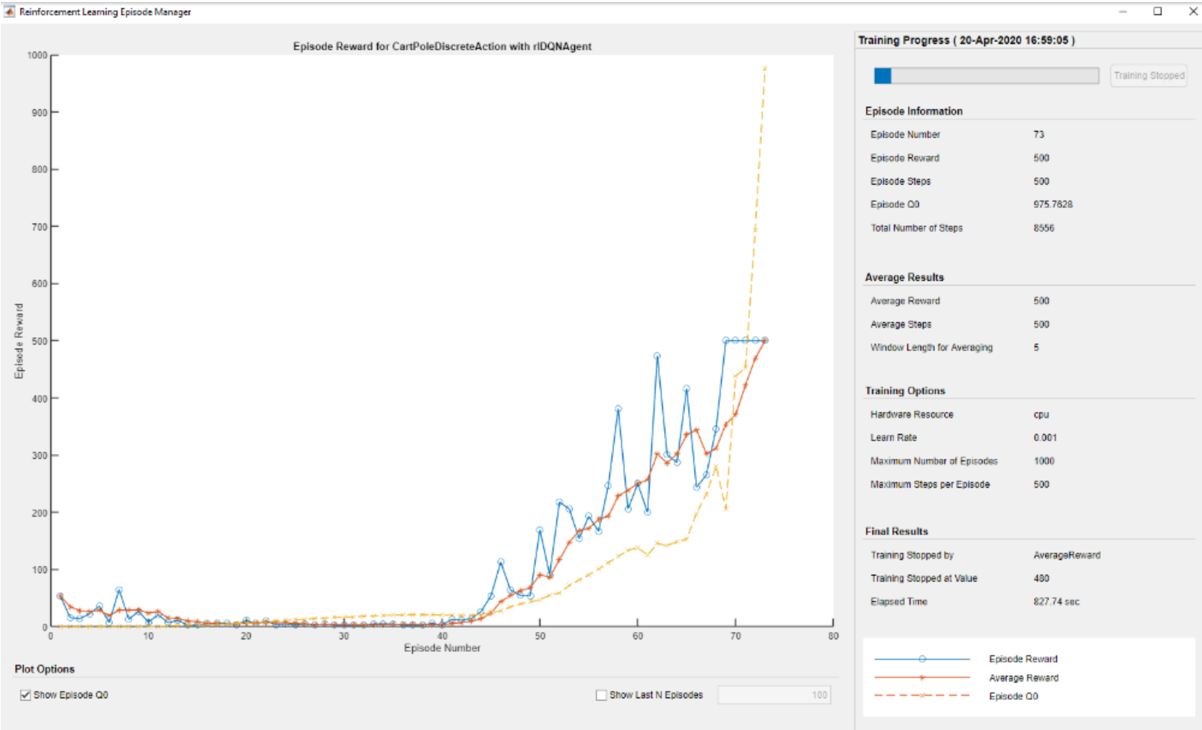


Figure 5: Training the DQN Agent

2.) <u>Pseudo-Code for a simple energy management system of a single battery:</u>

State = (SOC, P_{bat}, P_{fc})

Actions = (P_{bat}, del_P_{fc})

Update of the states:

SOC(t+1) = SOC–int (P_{bat})

P_{fc}(t+1) = P_{fc}(t) + del_P_{fc}

%Reward Logic

SOC > 35 and P_{load} – P_{bat} == 0

Reward = high

Else Reward = penalty

End

Similar to these two given examples, the feature of Reinforcement Learning Toolbox in Matlab can be used to create and simulate the complete energy management strategy for our project.

Data Driven Energy Management Strategy for Hybrid Electric Vehicles

As mentioned earlier, we have used Reinforcement Learning (RL) for energy management of HEV. In RL, learning is performed through feedback, which provides suitable compensation for the learning outcomes. The distinction between supervised learning and RL is that, unlike supervised learning, RL focuses on online performance, which is one of the benefits that it is more suitable for applications in real-time control techniques for HEVs. Q-learning is one of the RL algorithms used in this research.

Q learning is a method that allows the learning of optimal control online, where the Q function is learned using the temporal difference method based on interactions between the controller and environment. A novel energy management technique has been created specifically for the optimal power distribution problem of HEV control, based on Q-learning, as described in the introduction. The optimal control problem is first explained, then the new energy management technique is shown.

A. Optimal Control Problem:

The optimal control problem is defined to minimize the expected total cost over an infinite horizon as shown below:

$$\min \quad J_\pi(x_0) = \lim_{N\to\infty} \underset{d_k}{\mathrm{E}} \left\{ \sum_{k=0}^{N-1} \gamma^k g\big(x_k, \pi(x_k)\big) \right\} \qquad \textbf{(7)}$$

constrained by the following set of conditions:

$$\begin{gathered} \omega_{eng,min} \le \omega_{eng}(k) \le \omega_{eng,max} \\ T_{eng,min}(\omega_{eng}(k)) \le T_{eng}(k) \le T_{eng,max}(\omega_{eng}(k)) \\ T_{mot,min}(\omega_m(k), SOC(k)) \le T_{mot}(k) \le T_{mot,max}(\omega_m(k), SOC(k)) \\ SOC_{min} \le SOC(k) \le SOC_{max} \end{gathered} \qquad \textbf{(8)}$$

Here, x_k is the state variable, g is the instantaneous cost incurred, γ is the discount factor that represents the future cost as the expected value of the cost at current time step, $J_\pi(x_o)$ is the expected cost when the system starts at state x_o and follows the policy π and u is the engine power P_e.

The state variable x_k is composed of a four dimensional state space as given below:

$$x_k = [SOC, P_{dem}, v, E_{on}] \quad \textbf{(9)}$$

Here, SOC is the battery state of the charge, and E_{on} is the engine on/off state. The engine on/off state is considered to avoid fuel consumption due to frequent engine changes to the on/off states.

The instantaneous cost incurred g is defined as the equation below:

$$g = W_{fuel} + \zeta(SOC) + \beta \cdot \Delta E_{on} \quad \textbf{(10)}$$

Here, W_{fuel} is the instantaneous fuel consumption and β is the coefficient for the engine on/off penalty. $\zeta(SOC)$ is a term that penalizes the SOC deviation for charge sustenance as given below.

$$\zeta(SOC) = \begin{cases} \mu \cdot \left(SOC - SOC_{ref}\right)^2 & if\ SOC > SOC_{min} \\ C_{Penalty} & if\ SOC \leq SOC_{min} \end{cases} \quad \textbf{(11)}$$

Here, μ and $C_{Penalty}$ are positive constant values for the SOC deviation. The underlying meaning of the optimal control problem is that the overall expectation of the cost for the infinite horizon is minimized instead of for a finite horizon, therefore the control policy result is time invariant, which can be easily implemented as a real-time vehicle controller.

B. Model-Based Q Learning:

In Q-learning, the optimal cost $J^*(x_k)$ and optimal control policy $\pi^*(x_k)$ can be found as in the below equation using the Q-function:

$$\begin{aligned} J^*(x_k) &= \min_u (Q^*(x_k, u)) \\ \pi^*(x_k) &= \arg\min_u (Q^*(x_k, u)) \end{aligned} \quad \textbf{(12)}$$

Further, the Q Function can be updated as follows:

$$\begin{aligned} &Q(x_k, u_k) \\ &\leftarrow Q(x_k, u_k) + \alpha\left(g_k + \gamma \min_u Q(x_{k+1}, u) - Q(x_k, u_k)\right) \end{aligned} \quad \textbf{(13)}$$

When the system is in some state x_k , (i.e., in this HEV control problem, when the vehicle is in some state according to SOC_k , $P_{dem,k}$, v_k , and $E_{on,k}$), the control u_k is selected which has a minimum Q value. According to the action u_k, the state x_k changes to x_{k+1} with immediate reward g_k, then based on the Q value at the new state x_{k+1} and g_k, the Q function value $Q(x_k, u_k)$ is updated with the Bellman equation.

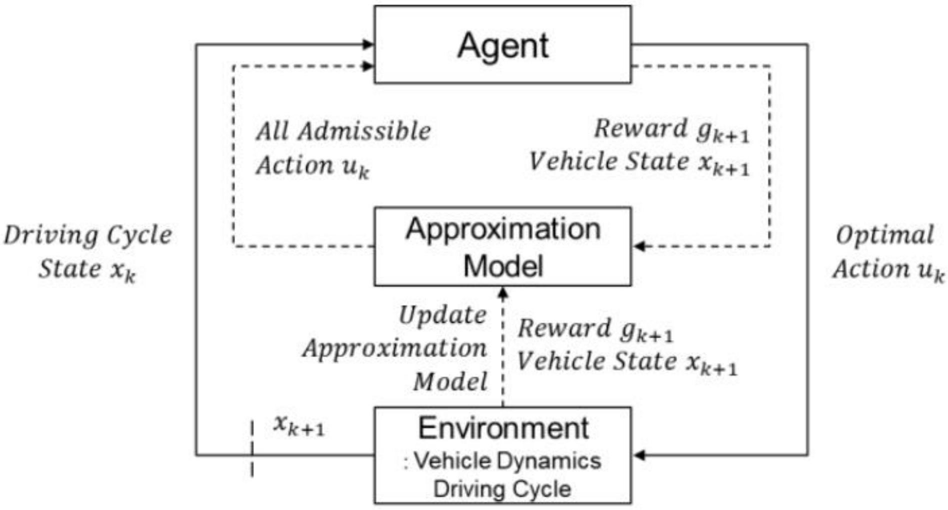


Figure 6: Concept of model-based reinforcement learning for an energy management strategy of hybrid electric vehicles

Pseudo Code of Algorithm for HEV Control:

Initialize $Q(x_k, u_k)$

Repeat each step $k = 1,2,3, \ldots$

1. Choose action optimal u_k based on $Q(x_k, u_k)$

2. Taking action u_k, observe reward $g(x_k, u_k)$, state x_{k+1}

3.1 Update model based on observation

$$g(x_k, u_k) \leftarrow g(x_k, u_k) + \alpha\big(g_k - g(x_k, u_k)\big)$$
$$x_{k+1}(x_k, u_k) \leftarrow x_{k+1}(x_k, u_k) + \alpha\big(x_{k+1} - x_{k+1}(x_k, u_k)\big)$$

3.2 Update Q using model for all admissible action u_k

$$Q(x_k, u_k) \leftarrow Q(x_k, u_k) + \alpha\left(g_k + \gamma \min_u Q(x_{k+1}, u) - Q(x_k, u_k)\right)$$

4. $x_k \leftarrow x_{k+1}$

Conclusion and Future Work

The project deals with energy management strategy of fuel-cell electric vehicle. Model Based Reinforcement Learning is used as the energy management strategy in FCEVs. The transition probability of the vehicle's driving speed profile is learned online based on driving data in the proposed RL based control strategy, and the control strategy is optimised using model-based Q-learning. In order to improve HEV fuel economy, it is required to not only improve the vehicle's engine efficiency, but also to define the vehicle's speed profile for use in the management strategy. The proposed control strategies in this project report have a powerful mathematical framework that uses reinforcement learning to model the driving cycle information from a stochastic perspective, followed by optimization using model-based approaches with an explainable and tunable vehicle state approximation model to solve the HEV supervisory control problem.

Experimental validation of the proposed control approach is required as future study. Because the control strategy is tested using simulations, experiments are required to confirm the approach. Furthermore, based on experimental evidence, the tradeoff relationship between computational burden and fuel economy performance of the technique should be examined. Finally, we expect that by combining the proposed technique with other practical challenges such as pollution or drivability, we will be able to make it more practical and feasible.

References

1.) Zhou, Y. F., Huang, L. J., Sun, X. X., Li, L. H., & Lian, J. (2020). A Long-term Energy Management Strategy for Fuel Cell Electric Vehicles Using Reinforcement Learning. *Fuel Cells*, *20*(6), 753–761. https://doi.org/10.1002/fuce.202000095

2.) Sun, H., Fu, Z., Tao, F., Zhu, L., & Si, P. (2020). Data-driven reinforcement-learning-based hierarchical energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles. *Journal of Power Sources*, *455*, 227964. https://doi.org/10.1016/j.jpowsour.2020.227964

3.) Fares, D., Chedid, R., Panik, F., Karaki, S., & Jabr, R. (2015). Dynamic programming technique for optimizing fuel cell hybrid vehicles. *International Journal of Hydrogen Energy*, *40*(24), 7777–7790. https://doi.org/10.1016/j.ijhydene.2014.12.120

4.) Panday, A., & Bansal, H. O. (2014). A Review of Optimal Energy Management Strategies for Hybrid Electric Vehicle. *International Journal of Vehicular Technology*, *2014*, 1–19. https://doi.org/10.1155/2014/160510

5.) Sulaiman, N., Hannan, M., Mohamed, A., Majlan, E., & Wan Daud, W. (2015). A review on energy management system for fuel cell hybrid electric vehicle: Issues and challenges. *Renewable and Sustainable Energy Reviews*, *52*, 802–814. https://doi.org/10.1016/j.rser.2015.07.132

6.) Lee, H., Kang, C., Park, Y. I., Kim, N., & Cha, S. W. (2020). Online Data-Driven Energy Management of a Hybrid Electric Vehicle Using Model-Based Q-Learning. *IEEE Access*, *8*, 84444–84454. https://doi.org/10.1109/access.2020.2992062

7.) Panday, A., & Bansal, H. O. (2016). Energy management strategy for hybrid electric vehicles using genetic algorithm. *Journal of Renewable and Sustainable Energy*, *8*(1), 015701. https://doi.org/10.1063/1.4938552

8.) Lee, H., Kim, N., & Cha, S. W. (2020). Model-Based Reinforcement Learning for Eco-Driving Control of Electric Vehicles. *IEEE Access*, *8*, 202886–202896. https://doi.org/10.1109/access.2020.3036719

9.) Lee, H., & Cha, S. W. (2021). Reinforcement Learning Based on Equivalent Consumption Minimization Strategy for Optimal Control of Hybrid Electric Vehicles. *IEEE Access*, *9*, 860–871. https://doi.org/10.1109/access.2020.3047497

10.) Singh, K. V., Bansal, H. O., & Singh, D. (2019). A comprehensive review on hybrid electric vehicles: architectures and components. *Journal of Modern Transportation*, *27*(2), 77–107. https://doi.org/10.1007/s40534-019-0184-3

11.) Lee, H., & Cha, S. W. (2021a). Energy Management Strategy of Fuel Cell Electric Vehicles Using Model-Based Reinforcement Learning With Data-Driven Model Update. *IEEE Access*, *9*, 59244–59254. https://doi.org/10.1109/access.2021.3072903