

Machine Learning-Based Energy Management in a Hybrid Electric Vehicle for better optimization of power

¹ Atul Dwivedi [1 dwivedi.20@gmail.com](mailto:dwivedi.20@gmail.com) ² Balendu Bhushan Pandey [2 blissbalendu@gmail.com](mailto:blissbalendu@gmail.com)

Abstract:

This paper investigates the energy management problem in hybrid electric vehicles (HEVs) focusing on the minimization of the operating cost of an HEV, including both fuel and battery replacement cost. More precisely, the paper presents a nested learning framework in which both the optimal actions (which include the gear ratio selection and the use of internal combustion engine versus the electric motor to drive the vehicle) and limits on the range of the state-of-charge of the battery are learned on the fly. The inner-loop learning process is the key to minimization of the fuel usage whereas the outerloop learning process is critical to minimization of the amortized battery replacement cost. Experimental results demonstrate a maximum of 48% operating cost reduction by the proposed HEV energy management policy.

I. Introduction:

Electric vehicles (EVs) and hybrid electric vehicles (HEVs) have been gaining market share nowadays in the automotive market due to the concerns about large amounts of fuel consumption and pollution resulted from the conventional internal combustion engine (ICE) vehicles. By integrating electric motors (EMs) into the vehicle propulsion system, EVs and HEVs achieve higher energy efficiency and lower pollution emission compared with the conventional vehicles. HEVs, which represent a transition from conventional ICE vehicles to full electric vehicles, have higher fuel efficiency than conventional vehicles and fewer battery-related problems than EVs. However, due to the hybrid structure of the propulsion system, advanced HEV energy management techniques are needed to fully explore the advantages of HEVs. The hybrid propulsion system of an HEV consists of an ICE and one or more EMs. The ICE converts chemical energy of the fuel into mechanical energy to propel the vehicle. The EM converts electrical energy stored in the battery pack to propel the vehicle, and it can also operate as a generator collecting kinetic energy during braking to charge the battery pack, which is called the regenerative braking, a mechanism improving the energy efficiency of EVs and HEVs. HEV energy management techniques coordinate the operation of ICE and EM to improve the energy efficiency of HEVs.

The fuel cost is one major operating cost component of the HEV. Therefore, the majority of previous work on HEV energy management aimed at improving the fuel economy. The rule-based strategies for HEV energy management interpret the driver controlled pedal motion into the required propulsion power, and determine the power split between the ICE and the EM based on intuition, human expertise or fuzzy logic. The optimization-based control strategies either minimize the fuel consumption during a trip with given, predicted or stochastic future driving profile, or perform control by converting battery charge into equivalent fuel consumption (ECMS and adaptive-ECMS approaches).

The state-of-health (SoH) of the HEV battery pack is degrading with the operation of an HEV due to the frequent charging/discharging of the battery pack by the EM. The work studied the SoH degradation model for the EV/HEV battery pack as a function of the state-of-charge (SoC) swing, the number of charging/discharging cycles, etc. The battery pack will reach its end-of-life when its SoH degrades to 80% or 70% and the battery pack replacement will result in additional operating cost of the HEV. Enlarging the battery pack energy capacity within size, weight and cost constraints is preferred for exploring the energy storage capability of the battery pack to improve the HEV fuel economy, and especially the plug-in HEV (PHEV) employs a higher-capacity battery pack which is charged using the grid power. The battery replacement cost increases significantly with enlarged battery capacity, and therefore the amortized battery replacement cost must not be neglected in the HEV. There are some work taking into account battery SoH degradation when optimizing the fuel efficiency. However, these work have one or more of the following shortcomings: (i) The HEV energy management policies they use are based on ECMS or adaptive ECMS approaches which rely on the knowledge of the future driving profile. If the prediction of the future driving profile is not accurate, the effectiveness of these ECMS and adaptive-ECMS based approaches can be degraded. (ii) They do not use accurate analytical battery SoH degradation model in the optimization and evaluation, instead, they use Ah-throughput or battery output power as the equivalent of the battery SoH degradation during charging and discharging processes.

Machine learning provides a powerful tool for the agent (i.e., decision-maker) to “learn” how to “act” optimally when the explicit and accurate system modeling is difficult or even impossible to obtain. The agent can observe the environment’s state and take an action according to the observed state. A reward will be given to the agent as a result of the action taken. Stimulated by the reward, the agent aims to derive a policy, which is a mapping from each possible state to an action, by “learning” from its past experience. The reinforcement learning has been applied to the HEV energy management problem, such that the HEV energy management policy does not rely on any knowledge of the future driving profile. An inverse reinforcement learning technique has been applied for learning the driver behavior, however, it is out of our focus. In this proposed work, we investigate the HEV energy management problem focusing on the minimization of the operating cost of an HEV, including both fuel cost and amortized battery replacement cost (i.e., battery purchase plus installation cost). We present a nested learning framework in which both the optimal actions (which include the gear ratio selection and the use of ICE versus EM to drive the vehicle) and limits on the range of battery SoC are learned on the fly. More precisely, the inner-loop learning process determines the operation modes of the HEV components whereas the outer-loop learning process modulates the battery SoH degradation from a global point of view. Due to the usage of the machine learning techniques, the proposed HEV energy management does not rely on perfect and accurate system modeling (i.e., HEV component modeling and driving profile modeling.) The proposed nested learning framework for HEV energy management differs from the reinforcement learning-based framework in that (i) the amortized battery replacement cost is incorporated into the HEV energy management; and (ii) two nested learning processes are used in which the inner-loop learning process is the key to minimization of the fuel usage and the outer-loop learning process is critical to minimization of the amortized battery replacement cost. Experimental results demonstrate a maximum of 48% operating cost reduction by the proposed HEV management policy.

II. System Description:

Although this work aims to design a smart HEV controller that discovers the energy management policy by learning from its experience, it is still necessary to understand the fundamentals of HEV operation. By way of an example and without loss of generality, we discuss the parallel HEV configuration as in most of the literature work on HEV energy management. There are five operation modes of a parallel HEV, depending on the flow of energy: (i) only the ICE propels the vehicle, (ii) only the EM propels the vehicle, (iii) the ICE and EM propel the vehicle in parallel, (iv) the ICE propels the vehicle and at the same time drives the EM to charge the battery pack, and (v) the EM charges the battery pack when the vehicle is braking (i.e., regenerative braking mode.)

HEV Component Analysis :

1) Internal Combustion Engine (ICE):

According to the quasi-static ICE model, the fuel efficiency of an ICE is calculated as $\eta_{ICE}(T_{ICE}, \omega_{ICE}) = T_{ICE} \cdot \omega_{ICE} / (m_f \cdot D_f)$. (1) In (1), T_{ICE} and ω_{ICE} are the torque (in N·m) and speed (in rad/s) of the ICE, respectively, which represent the operation point of the ICE. m_f is the fuel consumption rate (in g/s) of the ICE, depending on the ICE operation point. And D_f is the fuel energy density (in J/g). Figure 1 (a) represents the contour map of the fuel consumption rate of an example ICE in the ICE speed-torque plane.

2) Electric Motor (EM):

The EM operates either as a motor to propel the vehicle or as a generator to charge the battery pack. The efficiency of the EM is $\eta_{EM}(T_{EM}, \omega_{EM}) = (T_{EM} \cdot \omega_{EM}) / P_{batt}$ $T_{EM} \geq 0$ $P_{batt} / (T_{EM} \cdot \omega_{EM})$ $T_{EM} < 0$ where T_{EM} and ω_{EM} are respectively the torque and speed of the EM, and P_{batt} is the output power of the battery pack. When the EM operates as a motor, T_{EM} is positive and the battery pack is discharging i.e., $P_{batt} > 0$; when the EM operates as a generator, T_{EM} is negative and the battery pack is charging i.e., $P_{batt} < 0$. Figure 2 represents the efficiency contour map of the EM as a motor or a generator. To ensure safe and smooth operation of an EM, the following constraints should be satisfied: $0 \leq \omega_{EM} \leq \omega_{max EM}$, (4) $T_{min EM}(\omega_{EM}) \leq T_{EM} \leq T_{max EM}(\omega_{EM})$.

3) Vehicle Tractive Force:

The vehicle tractive force F_{TR} to support the vehicle speed and acceleration (which are determined by the driver through pressing the braking or acceleration pedal) is derived by $F_{TR} = m \cdot a + F_g + F_R + F_{AD}$ $F_g = m \cdot g \cdot \sin\theta$ $F_R = m \cdot g \cdot \cos\theta \cdot CR$ $F_{AD} = 0.5 \cdot \rho \cdot CD \cdot AF \cdot v^2$ where m is the vehicle mass, a is the vehicle acceleration, F_g is the force due to road slope, F_R is the rolling friction force, F_{AD} is the air drag force, θ is the road slope angle, CR is the rolling friction coefficient, ρ is the air density, CD is the air drag coefficient, AF is the vehicle frontal area, and v is the vehicle speed. Given v , a and θ , the tractive force F_{TR} can be derived using. Then, the vehicle wheel torque T_{wh} and wheel speed ω_{wh} are related to F_{TR} , v , and wheel radius r_{wh} by $T_{wh} = F_{TR} \cdot r_{wh}$, $\omega_{wh} = v / r_{wh}$. The demanded power for propelling the vehicle i.e., p_{dem} satisfies $p_{dem} = F_{TR} \cdot v = T_{wh} \cdot \omega_{wh}$.

III. A Nested Learning Framework For HEV Energy Management:

In this work, we aim to minimize the operating cost of an HEV including both fuel cost and amortized battery replacement cost. To achieve this goal, we propose a nested learning framework for HEV energy management, in which the optimal actions to propel the vehicle and the limits on the change in the SoC of the battery are learned on the fly by the inner-loop reinforcement learning and the outer-loop adaptive learning, respectively. The inner-loop reinforcement learning process is the key to minimization of the fuel usage, whereas the outer-loop adaptive learning process is critical to minimization of the amortized battery replacement cost.

A. Motivation :

We use reinforcement learning in the inner loop due to the following reasons. (i) The inner-loop HEV energy management aims to minimize the total fuel consumption during a driving trip rather than the instantaneous fuel consumption rate at each time step; the reinforcement learning also aims to optimize an expected cumulative return rather than an immediate reward. (ii) During a driving trip, the changes of vehicle speed, power demand and battery charge level require different HEV operation modes; the reinforcement learning agent takes different actions depending on the current state. (iii) The inner-loop HEV energy management does not have a priori knowledge of a whole driving trip, while it has only the knowledge of the current vehicle speed and power demand values and the current fuel consumption rate as a result of an action taken previously; the reinforcement learning agent only needs the knowledge of the current state and the current reward in order to learn the optimal policy, while it does not need to have knowledge of the system input in prior or the detailed system modeling. The inner loop is the key to minimization of the fuel usage, however, we also consider battery SoH degradation in the inner loop by incorporating the battery capacity fading term into the reward of the reinforcement learning, such that the inner loop itself can be used as an independent HEV energy management framework for minimizing the total operating cost

In the previous work on HEV energy management, a fixed battery SoC range is used i.e., the battery pack SoC is clamped by fixed lower bound and upper bound. Then, the resultant HEV energy management strategies may tend to use up the available battery energy to improve fuel economy even for some very short urban trips, which may harm battery SoH seriously. The battery can obtain significant amount of energy by regenerative braking in urban trips. It is not always necessary to use up the available battery energy.

B. Inner-Loop Reinforcement Learning Process:

1) Reinforcement:

Learning Background: In reinforcement learning, the decision-maker is called the agent and everything outside the agent is called the environment. At each of a sequence of discrete time steps $t = 0, 1, 2, \dots$. At each time step t , the agent observes the environment's state $s_t \in S$ and on that basis takes an action $a_t \in A$, where S and A are the sets of possible states and actions, respectively. One time step later, in part as a consequence of the action taken, the agent receives a numerical reward r_{t+1} and finds the environment in a new state s_{t+1}

A policy π of the agent is a mapping from each state $s \in S$ to an action $a \in A$ that specifies the action $a = \pi(s)$ that the agent will choose when the environment is in state s . The ultimate goal of an agent is to find the optimal policy, such that $V \pi (s) = E (X_{\infty} = \sum_{k=0}^{\infty} \gamma^k \cdot r_{t+k+1} \mid s_t = s)$ is maximized for each state $s \in S$. The value function $V \pi (s)$ is the expected return when the environment starts in state s at time step t and follows policy π thereafter. γ is a parameter, $0 < \gamma < 1$, called the discount rate that ensures the infinite sum (i.e., $\sum_{k=0}^{\infty} \gamma^k \cdot r_{t+k+1}$) converges to a finite value. More importantly, γ reflects the uncertainty in the future. r_{t+k+1} is the reward received at time step $t+k+1$.

2) State Space:

We define the state space of the inner-loop reinforcement learning where p_{dem} is the power demand for propelling the HEV, v is the vehicle speed, and q is the charge stored in the battery pack. Different actions should be taken under different states. For example, if the power demand is negative i.e., the vehicle is braking, the action taken by the agent (i.e., HEV controller) should be charging the battery by using the EM as a generator. On the other hand, if the power demand is a very large positive value, the action should be discharging battery to power the EM, which propels the vehicle in assistance with ICE.

$$S = \{s = [p_{dem}, v, q]^T \mid p_{dem} \in P_{dem}, v \in V, q \in Q\},$$

A reinforcement learning agent should be able to observe a state. In the actual implementation of the inner-loop reinforcement learning, the current power demand level p_{dem} and vehicle speed level v can be obtained by using sensors to measure the driver controlled pedal motion. However, the charge level q cannot be obtained from online measurement of terminal voltage, since the battery pack terminal voltage changes with the charging/discharging current and therefore it cannot be an accurate indicator of q . To observe the charge level q , the Coulomb counting method is needed by the agent, which is typically realized using a dedicated circuit.

P_{dem} , V , and Q are respectively the finite sets of power demand levels, vehicle speed levels, and levels of charge stored in the battery pack. Discretization is required when defining these finite sets. In particular, Q is defined by discretizing the range of charge stored in the battery pack i.e., $[q_{min}, q_{max}]$ into a finite number of charge levels: $Q = \{q_1, q_2, \dots, q_N\}$,

where $q_{min} = q_1 < q_2 < \dots < q_N = q_{max}$. Generally, q_{min} and q_{max} are 40% and 80% of the battery pack nominal capacity, respectively, in the charge-sustaining energy management for ordinary HEVs; 0% and 80%, respectively, in the charge-depleting energy management for PHEVs. In the outerloop adaptive learning process, we will optimize q_{min} value to modulate the battery SoH degradation and q_{max} is usually fixed in the HEV control. 3) Action Space: We define the action space of the innerloop reinforcement learning as a finite number of actions, each represented by the discharging current of the battery pack and the gear ratio value:

$$A = \{a = [i, R(k)]^T \mid i \in I, R(k) \in R\}$$

where an action $a = [i, R(k)]^T$ taken by the agent is to discharge the battery pack with current i and choose the k -th gear ratio. The set I contains within it a finite number of current values in the range of $[-I_{max}, I_{max}]$. Please note that $i > 0$ denotes discharging the battery pack; and $i < 0$ denotes charging the battery pack. The set R contains the allowable gear ratio values, which depend on the drivetrain design. Usually, there are four or five gear ratio values in total.

Alternatively, we can define a reduced action space A_{re} , in which an action $a_{re} = [i]$ is to discharge the battery pack with current i (and the gear ratio $R(k)$ is selected by solving an optimization problem such that the resultant fuel consumption rate is minimized.) The complexity and convergence speed of reinforcement learning algorithms are proportional to the number of state-action pairs. Therefore, the reduced action space A_{re} helps to reduce the complexity and increase convergence speed by a factor of four or five. However, this reduced action space relies on HEV component modeling when solving the optimization problem. In summary, we can either use the original action space for model-free control or use the reduced action space A_{re} for reduced complexity and increased convergence rate.

3) Reward:

The objective of the inner-loop reinforcement learning is to minimize the HEV operating cost including both fuel cost and amortized battery replacement cost. Therefore, we define the reward r that the agent receives after taking action a in state s as the negative of the weighted sum of the fuel consumption and battery capacity fading in that time step i.e., $-m' f \cdot \Delta T - w \cdot \Delta C_{f ade}$, where ΔT is the length of a time step, w is the weight of battery capacity fading (w is determined by the ratio of the fuel cost to the amortized battery cost), and $m' f$ and $\Delta C_{f ade}$ are respectively the fuel consumption rate and battery capacity fading in that time step. The reinforcement learning agent aims to maximize the expected return J , which is a discounted sum of rewards. Therefore, by using the negative of the weighed sum of the fuel consumption and battery capacity fading in a time step as the reward, the fuel consumption and battery capacity fading will be minimized while maximizing the expected return

4) TD(λ)-Learning Algorithm:

We adopt the TD(λ)- learning algorithm for deriving the optimal policy of the inner-loop reinforcement learning, due to its relatively higher convergence rate and higher performance in non-Markovian environment. In this algorithm, a Q value, denoted by $Q(s,a)$, is associated with each state-action pair (s,a) , where a state s is represented by the power demand p_{dem} , the vehicle speed v , and the battery charge level q , and an action a is to discharge the battery with current i and choose the k -th gear ratio. The $Q(s,a)$ value approximates the expected discounted cumulative reward of taking action a in state s . The TD(λ)- learning algorithm is summarized as follows.

In the TD(λ)-learning algorithm, the Q values are initialized arbitrarily at first. At each time step t , the agent first selects an action a_t for the current state s_t based on the $Q(s,a)$ values. To avoid the risk of getting stuck in a sub-optimal solution, the exploration-exploitation policy is employed for the action selection, i.e., the agent does not always select the action a that results in the maximum $Q(s_t, a)$ value for the current state s_t . After taking the selected action a_t , the agent observes a new state s_{t+1} and receives reward r_{t+1} . Then, based on the observed s_{t+1} and r_{t+1} , the agent updates the $Q(s,a)$ values for all the state-action pairs, in which the eligibility $e(s,a)$ of each state-action pair is updated and utilized during the Q value update. The eligibility $e(s,a)$ of a state-action pair reflects the degree to which the particular state-action pair has been encountered in the recent past and λ is a constant between 0 and 1. Due to the use of the eligibility of the state-action pairs, we do not need to update Q values and eligibility e of all state-action pairs. We only keep a list of M most recent stateaction pairs since the eligibility of all other state-action pairs is at most λM , which is negligible when M is large enough.

5) Application Specific Improvement of the TD(λ):

Learning Algorithm: We modify the TD(λ)-learning algorithm to improve its performance and convergence rate in the HEV control scenario by accommodating different operation modes of an HEV. Specifically, when selecting an action for the current state, the agent takes into account the actual HEV operation mode besides the stored Q values. For example, if the power demand is negative i.e., the regenerative braking mode, the agent will definitely choose the maximum allowable charging current for the battery pack to harvest the kinetic energy as much as possible. If the battery charge level is very high, the agent will use EM power with higher likelihood to propel the vehicle. And if the battery charge level is very low, the agent is likely to use more ICE power to propel the vehicle and at the same time charge the battery. In summary, these application specific modifications significantly improve performance and convergence rate of TD(λ)-learning algorithm.

6) Complexity and Model-Free Analysis:

The time complexity of the TD(λ)-learning algorithm at a time step is $O(|A| + M)$, where $|A|$ is the total number of actions and M is the number of the most recent state-action pairs kept in memory. Usually, $|A| + M$ is within a few hundred, and therefore, the algorithm has negligible computation overhead. In terms of convergence rate, normally, the TD(λ)-learning algorithm can converge within L time steps, where L is approximately three to five times of the number of state-action pairs. In simulation, due to the application specific improvement, the TD(λ)-learning algorithm can converge within one hour driving, which is much shorter than the total lifespan of an HEV. To further speed up the convergence rate, the Q values can be initialized by the manufacturers with optimized values.

III. Experimental Results:

We simulate the operation of a PHEV, the model of which is developed in the vehicle simulator ADVISOR [1]. The key parameters of the PHEV are summarized in Table I. We test our proposed policy and compare with the reinforcement learning (RL) policy and the rule-based policy. We use both real-world and testing driving trip profiles, which are developed and provided by different organizations and projects such as U.S. EPA (Environmental Protection Agency) and E.U. MODEM (Modeling of Emissions and Fuel Consumption in Urban Areas project).

TABLE I. PHEV KEY PARAMETERS

Vehicle	Transmission	ICE
$m = 1254$ kg $CR = 0.009$ $CD = 0.335$ $AF = 2$ m ² ; $rwh = 0.282$ m	$\rho_{reg} = 1.75$ $\eta_{reg} = 0.98$ $\eta_{gb} = 0.98$ $R(k) = [13.5; 7.6$ $5.0; 3.8; 2.8]$	Peak power 41Kw peak eff. 34% EM peak power 56kW peak eff. 92%
	Battery Capacity 25A·h Voltage 240V	

Table II presents the simulation results of the operating cost of the PHEV during different driving trips when the proposed, the RL, and the rule-based policies are adopted. For example, as shown in Table II,

the proposed policy results in 0.0028% battery capacity fading and 344.17g fuel consumption in the MODEM5713 driving trip, which correspond to \$0.76 amortized battery replacement cost and \$0.37 fuel consumption cost, and the total operating cost is \$1.13. When calculating the operating cost, we use the America average gasoline price of \$3/gal and the total battery replacement cost of \$8,000 for the PHEV. Generally, the battery replacement cost of a PHEV is in the range \$10,000~\$12,000 [3] for battery pack with average capacity of 10kWh. We use the battery replacement cost of \$8,000 for the 6kWh battery. We use the complete cycle-decoupling method [30] to evaluate the battery capacity fading during each trip. From Table II we can observe that the proposed policy consistently achieves the lowest operating cost comparing with the RL and rule-based policies. The proposed policy achieves a maximum of 47% operating cost reduction comparing with the rule-based policy, and a maximum of 48% reduction comparing with the RL policy.

TABLE II. OPERATING COST OF THE PHEV IN DIFFERENT TRIPS USING THE PROPOSED, RL, AND RULE-BASED POLICIES.

Trip	Trip Proposed	RL	Rule
MODEM 5713 cost	0.0028%(\$0.76) +344.17g(\$0.37) =(\$1.13)	0.0045%(\$1.22) +310.56g(\$0.33) =(\$1.55)	0.0044%(\$1.18) +383.30g(\$0.41) =(\$1.59)
Hyzem Motorway cost	0.0018%(\$0.50) +1991.9g(\$2.16) =(\$2.66)	0 0.0048%(\$1.28) +2001.9g(\$2.17) =(\$3.45)	.0050%(\$1.36) +2093.6g(\$2.27) =(\$3.63)
FTP75 cost	0.0027%(\$0.73) +311.40g(\$0.33) =(\$1.06)	0.0043%(\$1.16) +295.97g(\$0.32) =(\$1.48)	0.0048%(\$1.30) +623.73g(\$0.67) =(\$1.97)
US06 cost	0.0028%(\$0.74) +414.17g(\$0.45) =(\$1.19)	0.0043%(\$1.17) +354.34g(\$0.38) =(\$1.55)	0.0036%(\$0.98) +321.02g(\$0.34) =(\$1.32)
UDDS cost	0.0032%(\$0.85) +298.48g(\$0.32) =(\$1.17)	0.0044%(\$1.19) +355.85g(\$0.38) =(\$1.57)	0.0048%(\$1.30) +630.22g(\$0.68) =(\$1.98)
OSCAR cost	0.0021%(\$0.57) +149.51g(\$0.16) =(\$0.73)	0.0043%(\$1.16) +222.75g(\$0.24) =(\$1.40)	0.0042%(\$1.12) +242.54g(\$0.26) =(\$1.38)

Based on Table II, we also have the following observations: (i) For a PHEV, the amortized battery replacement cost is a large portion of the total operating cost and is even higher than the fuel cost for some driving trips. (ii) The relative amortized battery replacement cost is more significant for shorter driving trip. (iii) Our proposed policy can prolong the battery life significantly besides reducing the operating cost. (iv) Although the RL policy can reduce the fuel consumption comparing with the rule-based policy, in some case the operating cost from the RL policy is even higher because the RL policy does not take into account the battery cost when optimizing the fuel consumption. (v) The amortized battery replacement cost is non-negligible when optimizing the total operating cost.

V. Conclusions:

This paper investigates the HEV energy management problem for the minimization of the operating cost of an HEV by using a nested learning framework. The inner loop determines the operation modes of the HEV components and is the key to minimization of the fuel usage, whereas the outer loop modulates the battery SoH degradation globally. Experimental results demonstrate a maximum of 48% operating cost reduction by the proposed HEV energy management policy.

References:

- [1] ADVISOR 2003 documentation. National Renewable Energy Lab.
- [2] High-performance battery monitor IC with coulomb counter, voltage and, temperature measurement. Texas Instruments.
- [3] [http://batteryuniversity.com/learn/article/hybrid electric vehicle](http://batteryuniversity.com/learn/article/hybrid%20electric%20vehicle).
- [4] E. Alpaydin. Introduction to machine learning. MIT press, 2004.
- [5] H. Banvait and et al. A rule-based energy management strategy for plug-in hybrid electric vehicle (phev). In ACC'09.
- [6] A. G. Barto. Reinforcement learning: An introduction. MIT press, 1998.
- [7] B. M. Baumann and et al. Mechatronic design and control of hybrid electric vehicles. Mechatronics, IEEE/ASME Trans, 2000.
- [8] H. Borhan and et al. Mpc-based energy management of a power-split hybrid electric vehicle. Control Systems Technology, IEEE Trans, 2012.
- [9] C. Chan. The state of the art of electric, hybrid, and fuel cell vehicles. Proceedings of the IEEE, 2007. [10] C.-C. Chan and et al. Electric, hybrid, and fuel-cell vehicles: Architectures and modeling. Vehicular Technology, IEEE Trans, 2010.
- [11] S. Delprat and et al. Optimal control of a parallel powertrain: from global optimization to real time control strategy. In Vehicular Technology Conference, 2002.
- [12] S. Ebbesen, P. Elbert, and L. Guzzella. Battery state-of-health perceptive energy management for hybrid electric vehicles. Vehicular Technology, IEEE Transactions on, 61(7):2893–2900, 2012.