

Data Driven Control

Opportunities, Benefits and Challenges



Abstract

With the extensive adoption of state-of-the-art sensing, communication and information management technologies in industries, large volumes of data corresponding to the different variables in complex industrial processes and engineering systems is available in many application areas. This data can be used to design more efficient, reliable, robust and adaptive control systems.

Data can be good representative of the inherent nonlinearity and stochasticity in a system. When a system gets too complex to accurately model mathematically, data can be utilized to recognize the behavior of these system and design reliable controllers for them. For example, a HVAC system of a large building is a complex system having more than 40 equations despite making simplifying assumptions about the Physics. Such a system of equations would be computationally difficult to solve and moreover the solution will not be robust and reliable as the simplifying assumptions render the model unable to take into consideration all the extreme case scenarios.

Hybrid approaches can also be used when in one subsystem, the data is not available in sufficient quantity due to physical sensing limitations and in another subsystem, the data is available aplenty but the subsystem is too complex to describe in closed form mathematical expressions. For example, a system where the physics model of one subsystem is well tested and reliable but the data collection is difficult such as a combustion chamber in a thermal power plant and another system where the stochasticity are not fully identified and accounted where data can be of utility to model the behavior of the system.

The two important factors in setting up a Data Driven Control (DDC) are domain knowledge and data availability – the mathematical functional structure of the controller model and the availability or unavailability of optimal actuation/decision commands for input instances. The domain knowledge is helpful in incorporating useful inductive bias in the model that is the structure in which the data needs to be fed and the a priori knowledge regarding what the useful features might be. The amount of data in terms of variance and volume is important to capture all the information concerning the behavior of the system. Based on these two factors a DDC problem can be set up in one of these four Machine Learning settings: Parameter Estimation of Fixed Structure Models, Supervised Learning, Unsupervised Learning and Reinforcement Learning can be used for Data Driven Control.

Reinforcement Learning (RL) is the most generalizable framework to model and solve complex control problems as it only needs a final goal which the model is to be optimized for. RL has provision for unbiased exploration when searching in solution space allowing the discovery of novel and unintuitive but optimal or near optimal control strategies. It can be used to learn from simulations. So, actual field data from an environment or a simulated model of the less stochastic environments that has the action versus response mapping such as video games is required and the agent or the controller model can learn up any arbitrary nonlinearity as required to capture the input signal to output actuation mapping that tries to optimize a specific defined control criterion.

Siemens has a wide range of complex industrial process plants and engineering systems. Plus, the MindSphere platform has enabled collection of enormous amounts of process and field data. This data can be utilized to design more reliable and adaptive controllers to get desirable smooth as well as robust control behavior. RL can also be used for designing more efficient resource allocation systems for flow optimization, energy management and inventory management. We present potential use cases for DDC and RL at Siemens.

4-5	Introduction
4-4	Classical Control Theory
4-5	Modern Control Theory
5-5	Data Driven Control
5-5	Reinforcement Learning
6-7	Model Based Control to Data Driven Control
8-8	Data Driven Control – When/why to use?
9-9	Advantages of Data Driven Control
10-10	Challenges in Data Driven Control
11-11	Classification of Data Driven Control Approaches
11-11	Indirect vs Direct Methods
11-11	Iterative vs Noniterative Methods
11-11	Online vs Offline Methods
12-12	Machine Learning in Data Driven Control
12-12	Control Parameter Identification
12-12	Supervised Learning
12-12	Unsupervised Learning
12-12	Reinforcement Learning
13-13	Existing Approaches to Data Driven Control
14-14	Existing Attempts at Data Driven Control
15-15	Reinforcement Learning as an Approach to Data Driven Control
18-18	Sim2Real
17-17	Benefits of Data Driven Control
17-17	Engineering Benefits
17-17	Financial Benefits
17-17	Marketing Benefits
18-19	References
20-20	Abbreviations
20-20	Contact

Contents

The motivation behind this white paper is to present the need, advantages and challenges in data driven control. It is targeted at the classical/modern model based control community and the business units to highlight the opportunities Data Driven Control presents and how it can be leveraged to create more robust, adaptive and reliable control systems and to create more efficient resource allocation solutions. In this white paper, we first set the context for the need for data driven approaches in design of controllers in complex engineering systems. We then describe the various advantages and challenges in the development of Data Driven Control (DDC) systems. We describe the various methods for DDC and their classification. We also describe the engineering, financial and marketing benefits

of DDC. We present how the DDC problem can be addressed in various Machine Learning settings. We then articulate the benefits and challenges in use of Reinforcement Learning for DDC.

Introduction

More than 2300 years ago, Aristotle first described the concept and purpose of Control Systems as a means to help automate the tasks otherwise done by humans [1]. Although not explicitly mentioned, the precision required in automated systems was implied when he said "...if every instrument could accomplish its own work, obeying or anticipating the will of others..."

Control engineering is a discipline that deals with design of automatic controllers to achieve desired behaviors in the systems or processes being controlled. The control systems use sensors to measure the input to and most of the times the output (in feedback control) from the mechanism/ process – henceforth called 'plant' – being controlled. It uses the measured input and output values to determine the actuation to be applied.

Classical Control Theory

Classical control theory deals with the behaviors of dynamical systems with inputs and how their behavior can be modified with feedback [26]. A closed loop control system monitors the output and compares it with the desired reference value. This difference between actual and desired output called the error signal is applied as feedback to the input of the system to bring the actual output closer to the reference.

There are several advantages of closed-loop control over open-loop control. Disturbance rejection and more robust performance in the face of model uncertainties can be expected as the error itself is being fed back as input to finetune the output. This implies reduced sensitivity to parameter variations and improved performance w.r.t. reference tracking. Sometimes the closed-loop and open-loop controls are used together where the open-loop control called feedforward serves to further improve tracking performance.

The physical system modelled in time domain take the form of high order differential equations which are difficult to solve. Classical control uses the Laplace transform to convert these differential equations in time domain into a regular algebraic polynomial in the transform s-domain.

Modern Control Theory

Modern control theory, instead of changing domains to avoid the time domain ODE math, converts the differential equations into a system of lower order time-domain equations called state equations that can be solved using numerical linear algebra techniques [27].

A way to classify control systems is based on whether they use the mathematical model of the plant or the real/ simulated I/O data of the plant or both. Model based control techniques involve first building the plant model aka system identification and then designing the control system based on this model(Fig 1). Plant model is the physics model of the system/ process to be controlled. Controller model is the model mapping input signals to output actuations or decision variables. It makes use of the Certainty Equivalence principle which says that for discrete time centralized systems with only additive uncertainty, the optimal control solution obtained is same as that would be obtained in absence of additive disturbances [27]. The model-based control relies on the faith that plant model represents the true system.

Adaptive control, robust control and optimal control are some of the more sophisticated techniques in modern control theory [28]. An adaptive controller must be adapt to a control system with parameters which vary or are initially uncertain. Adaptive control differs from robust control in that it does not need a priori information about the bounds on these uncertain or time-varying parameters. Robust control guarantees that if the changes are within given bounds the control law need not be changed while adaptive control is concerned with control law changing itself [31]. Robust control policy is static and designed to work assuming that certain variable might be unknown but bounded within known limits.

Optimal control is concerned with finding a control law for a given system such that a certain optimality criterion is achieved [30]. It is an extension of variational calculus. Stochastic control assumes that random noise with known probability distribution affects the evolution and observation of the state variables [32]. Nonlinear control theory deals with systems that are nonlinear, time-variant or both [28]. In hierarchical control systems, a set of devices and

governing software is arranged in a hierarchical tree [28]. The lower layers have local tasks, goals, and sensations, and their activities are planned and coordinated by higher layers which do not generally override their decisions.

Intelligent or data-driven control includes many techniques namely – machine learning control, neural network control, reinforcement learning control, Bayesian control, fuzzy control, neuro-fuzzy control, expert systems and genetic control [29].

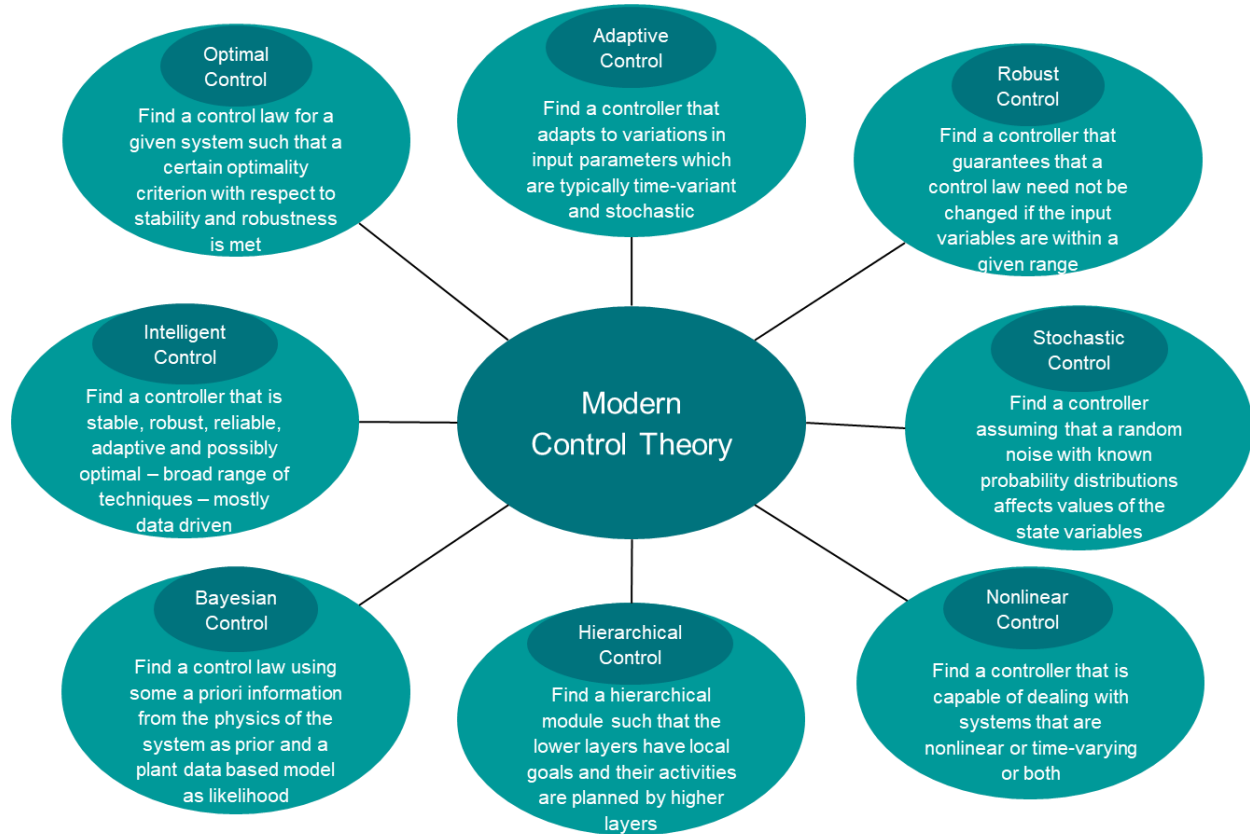


Fig 1: Branches of Modern Control Theory

Data Driven Control

In Data Driven Control systems, the identification of the process model and/or the design of the controller are based entirely on experimental or real-time field data collected from the plant [2][3]. When the systems get too complex to model mathematically, data can be an excellent alternative approach for identifying the model and/or designing the controller. The extensive adoption of state-of-the-art sensing, communication and information management technologies have enabled large volumes of data to be collected and stored for analysis.

This data can be used either indirectly for building the plant model first and then building the controller model on top of it or it can be directly used to build the controller model, skipping the need for a plant model altogether. Both offline data which is the historical experimental/ field data or the online data which is the continuous operational data can be used as it becomes available. Data can be used iteratively in batches or noniteratively all at once depending on the optimization (training) algorithm.

Reinforcement Learning

Reinforcement Learning is an area of Machine Learning that deals with the problem of finding optimal actions that software agents should take in an environment when in a particular state of inputs in order to maximize a long term cumulative reward or minimize a long term cumulative cost [4]. In a typical RL setting, the environment is modelled as a Markov Decision Process with special properties according to the problem at hand and the agent can be any general function approximator starting from a simple lookup table to complex deep neural networks. The agent

environment interactions of state-action-reward-state are simulated many times over in iterations or episodes and the agent learns through these interactions to take actions in a state that maximize the cumulative reward in the long term. Reinforcement Learning can be used to design optimal or near-optimal controller systems – it is also broadly referred to as a solution framework for the ‘Learning to Control’ problem.

Model Based Control to Data Driven Control

Modern model-based control techniques although successful in many cases have many drawbacks. Modelling, whether by first principles or from identification of data is always an approximation of the true system. Unmodelled system and unaccounted stochasticity arise from the simplifying assumptions made while modelling the system. Unmodelled complexities can arise from temporal variations as for some plants the parameters may vary quickly or their structure changes over time [2].

The answer to the question when are these approximations good enough, varies from system to system and can only reliably be evaluated empirically. In model-based control other variables such as poorly known parameters, choice of sensors, choice of actuators and interaction with environment must also be explicitly modelled and hence any model-based control system is not adaptable to variations in the operating conditions and must be remodeled again from scratch. Whereas the same data-based control system just needs to be finetuned with new data and would work well.

Take for example the controller for automotive differential that transmits the power from the transmission to the wheels and allows selective locking in the driveshaft to enable differential speeds in the two wheels while turning. Consider a scenario when the road surface is different under left and right wheels because of some water or oil spilled on the road. So, the changing coefficient of friction on the road requires different amount of torque to be applied to lock or unlock the differential. Such variations might be difficult to model mathematically whereas a data driven model might learn the adaptive torque function needed for stable cruise on varying road conditions.

Data can become the savior for complex nonlinear large-scale stochastic time-variant systems where accurate plant/process models do not exist. Data to the rescue – simulated or real field data from the plant can be used to:

- Finetune existing controller models
- Design the complete controller
- Predict and assess system rates
- Evaluate performance
- Make decisions regarding maintenance
- Diagnose fault events

One can design greedy or elementally locally optimal control solutions for complex systems but such controllers might not give global or full-scale optimal performance. Actual observed/ simulated (Monte Carlo) data in sufficient quantity and variation will always be better representative of system stochasticity than any assumed additive or multiplicative noise model. Hence leveraging the data while determining the control behavior can lead to more robust, more accurate and possibly globally optimal control solutions.

Data driven-control methods can be classified based on what kind of data was used for – online or offline data. Online data is the system I/O data within a finite time window. It reflects the current system state timely. The control system can capture and adapt to the variations if online data is used effectively. Offline data contains a great deal of information with respect to the system operation. Potential rules and patterns corresponding to system operations can be found through processing and mining and these temporal dependencies captures can be used to build smarter predictive controls. There are hybrid methods that make use of both offline and online data and try to capture benefits of both the approaches.

Some of the examples of Online Data Driven Control techniques include Simultaneous Perturbation and Stochastic Approximation (SPSA), Model Free Adaptive Control, Unfalsified Control Methodologies, etc. Some of the examples of Offline Data Driven Control techniques include PID Control, Iterative Feedback Tuning, Correlation based Tuning, Virtual Reference Feedback Tuning, Non-Iterative Data-Driven Model Reference Control, Sub-space approaches, etc.

Reinforcement Learning framework can be used as a much more general technique for data-driven control that can be used in multitude of problem settings: stochastic as well as deterministic, partially observable as well as fully observable, time-invariant as well as time-variant, continuous time as well as discrete event systems, multi-agent as well as single-agent environments, finite-time as well as infinite-time episodes, single-state as well as multi-state settings, continuous as well as discrete as well as categorical state-spaces, single-objective as well as multi-objective control problems.

Adaptive Control	Robust Control
<ul style="list-style-type: none"> • Must be adapt to a control system with parameters which vary or are initially uncertain • Does not need a priori information about the bounds on these uncertain or time-varying parameters • Dynamic and designed to work in a fashion to quickly adapt to the variations in the stochastic time-varying input parameters 	<p>Guarantees that if the changes are within given bounds the control law need not be changed while adaptive control is concerned with control law changing itself</p> <ul style="list-style-type: none"> • Does need a priori information about the bounds on these uncertain or time-varying parameters • Static and designed to work assuming that certain variable might be unknown but bounded within known limits

Data Driven Control – when/why to use?

- Many industrial processes, infrastructure systems, transportation systems, power management systems, large engineering products are becoming more and more complex [3]
- Large-scale and complex internal architectures of these systems make it difficult to create accurate plant models for them.
- Any simplifying assumptions made to create a model will bring in approximations and some phenomena in the system may be left unmodelled. This leaves the model unreliable.
- Due to the growing adoption state-of-the-art Sensing, Communication and Information Management technologies in industries, large volumes of data, widely called big data, are available – both in the form of large historical data from previous measurements and online data in real time during process runs.
- This data can be used to directly design controller, predict and assess system states, evaluate performance, make decisions, perform real time optimization and conduct fault diagnosis.

Advantages of Data Driven Control

- It is often very hard to get the closed form expression of some of the complex systems to be controlled – so it is difficult to get optimal control solution using conventional techniques [3].
- Greedy or elementally optimal control solutions for complex systems might not be optimal control solutions for the complete complex system.
- Actual observed/simulated (Monte Carlo) data accounts for/represents the actual stochasticity system faces in its operating environment.
- Thus, closed form equation models of such complex stochastic systems based on some simplifying assumptions do not account for all the realistic stochasticity the system faces whereas the data represents this stochasticity in a sampled form.
- Hence leveraging the data while determining the control behavior can lead to more efficient and more robust control policies.
- Data driven model is easier to update to changing operational conditions. The existing model just needs to be further trained and finetuned on the newly available data.

Take for example a Traffic Signal Control System – it might be possible to correctly model the traffic dynamics at a single signal but in real life scenarios there will always be multiple traffic signals in any area and the decisions of one agent affect that of all others. This interaction dynamics is mathematically complex and it is not possible to get closed form expression and the problem is nonconvex and high-dimensional. Hence data driven simulation based optimization is the best way to go for such a problem.

Challenges in Data Driven Control

- How much data is good enough to produce a reliable control model is an open question. Typically, empirical evaluation is conducted on a hold-out set often called 'test set' to assess the performance and reliability of the model.
- Very few theoretical studies are carried out on convergence, stability, robustness and optimality guarantees of data driven models. Due to the inherent nonlinearity and complexity in the systems, it is difficult to assert whether the convergent value of parameters obtained after many iterations is a global optimum or a local optimum [2].
- Highly complex data driven are typically treated as black-box models. They give good output decisions given some input but are not interpretable.
- The lack of interpretability leads further to lack of tractability and accountability within the model. For example, something going wrong somewhere in the system is not easy to track, address and resolve. The whole model might need to be retuned or replaced.

Classification of Data Driven Control Approaches

Indirect vs Direct Methods

Indirect methods take the standard two-step approach – first identifying the model and then tuning the controller based on such model. The main issue in this is that the controller is computed from the estimated model which itself is an approximation of the ground reality. To overcome this problems, direct methods map the experimental data directly on to the controller without any model to be identified in between [33].

Iterative vs Non-iterative Methods

In Iterative methods, repeated iterations are performed during the optimization problem with updates being performed as results of previous iteration. This approach is prone to online implementations. In noniterative methods, the optimal control parametrization is provided with a single optimization problem. This is important for cases where repetitions of data collection experiments are limited and the selected design technique should be capable of delivering a controller on a single data set. Noniterative methods are typically implemented offline [33].

Online vs Offline Methods

On many practical industrial applications where open-loop and closed-loop data are often available continuously, online DDC techniques use those data improve the quality of identified model and the performance of the controller each time new information is collected on the plant. Offline DDC work on batch of data which may be collected once or multiple regular long intervals of time [33]. (See Fig 2)

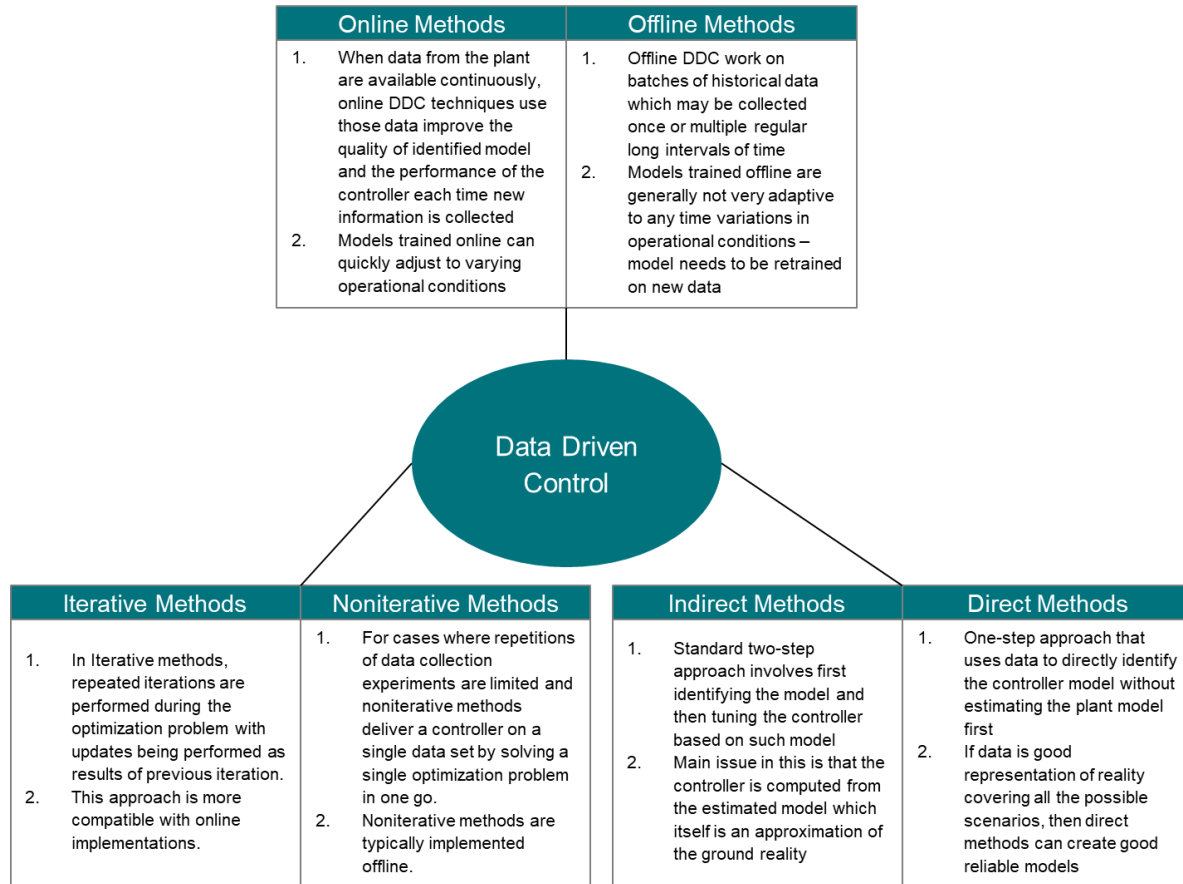


Fig 2: Classification of Data Driven Control Methods

Machine Learning in Data Driven Control

When learning from data, there are four aspects based on domain knowledge and data availability to a Machine Learning algorithm and the (to be) learned model:

1. The mathematical functional structure of the controller model
2. Final goal which the model is to be optimized for
3. Availability or unavailability of optimal actuation/ decision commands for input instances
4. Provision for unbiased exploration when searching in solution space

Based on these four aspects, four major categories of Data Driven Learning algorithms come up in the context of Control System Design.

Control Parameter Identification

- Structure of the control law is given but the parameters are unknown
- Examples: genetic algorithm for optimizing the coefficient of PID controller, discrete-time optimal control [34]

Supervised Learning Approach

- Control design as a regression problem of the first kind
- MLC to approximate a general nonlinear mapping from sensor signals to actuation commands
- Examples: Neural network as the function approximator for mapping input signals to actuation commands

Unsupervised Learning Approach

- Control design as a regression problem of the second kind
- MLC to identify arbitrary nonlinear control laws which minimize the cost function of the plant
- Examples: Optimization problem on control performance cost function as measured in the plant – using genetic programming as a regression technique [34]

Reinforcement Learning Approach

- Neither of the model, control law structure, optimizing actuation command need to be known
- Control law may be continually updated over measured performance changes
- Highly generalizable approach applicable in variety of settings

The more restrictions on the formulation and solution of the optimization problem, the more is the bias of a priori knowledge introduced in the model and intuitively lesser the data required to build a reliable model. RL introduces the least number of restrictions and the model hence is the most data-hungry method for DDC. (See Fig 3)

Requirement for DDC	Parameter Identification	Supervised Learning	Unsupervised Learning	Reinforcement Learning
1. Mathematical function structure of the controller model – Inductive Bias	Required	Sometimes assumed – neural networks can theoretically approximate any arbitrary nonlinear model	Assumed – Autoencoder Neural Networks can approximate any arbitrary low dimensional representation	Sometimes assumed – Neural networks as value function approximators can in theory approximate any nonlinearity
2. Availability of optimal actuation/decision commands for different input instances – Labelled data	Required	Required	Not Required	Not Required – certain techniques such as imitation learning can utilize this information
3. Smart exploration in search space – Search bias	Standard gradient based methods	Standard gradient based methods	Expectation Maximization, standard gradient based methods	Gradient based methods and stochastic monte carlo optimization
4. Some control performance which the model is optimized for – Final goal	Yes – minimize error as a function of difference between actual and predicted values of labels	Yes – minimize error as a function of difference between actual and predicted values of labels	Yes – minimize the variance within a group and maximize the variance between groups	Yes – maximize the final net cumulative reward or minimize the net cumulative cost

Fig 3: Spectrum of Machine Learning Methods in Data Driven Control based on Data Used

Existing Approaches to Data Driven Control

1. Simultaneous Perturbation Stochastic Approximation (SPSA) [5] is a DDC that uses online data. It uses only closed-loop measured data. It assumes that the nonlinear dynamics of the controller plant are unknown and the controller serves as a function approximator whose structure is fixed with tunable parameters. The approximator can be a neural network or a polynomial or any other type of approximator. Inputs to the neural network are the control signals and system outputs within a fixed time window before the current instant and the one step ahead desired output. The function approximator is trained to minimize a certain control performance index. The mathematical plant model is unknown hence the simultaneous perturbation stochastic approximation used instead of gradient descent.
However, there are some drawbacks to the SPSA method. Stochastic perturbation to the parameters may lead to wasted product – convergence rate is slow and not suitable for controlled plants whose parameters may vary quickly over time.
2. Model Free Adaptive Control [6] is another online DDC technique. It does not require a mathematical model of the controlled plant and assumes an equivalent dynamic linearization data model. It has a time varying incremental form with simple structure and less parameters. The main drawback of Model Free Adaptive Control is that it does not potentially have enough capacity to full complexity of the controller model such as nonlinearity and stochasticity.
3. Unfalsified control methodology [2] recursively falsifies control parameter sets that fail to satisfy a threshold performance criterion.
4. PID [8] control is a widely used DDC that uses offline data with a fixed model structure. It uses the information from proportional, integral and derivative of the error signal to finetune the model. Proportional component accounts for current error, integral term for the accumulated/ cumulative error and the derivative term accounts for the rate of change of error in finetuning the model.
5. Iterative feedback tuning [7] is a fixed structure offline DDC that does iterative optimization of the parameters of the fixed controller according the estimated gradient of a control performance criterion.
6. Correlation based tuning [2] is a fixed model structure offline DDC whose underlying idea inspired by the correlation approach in system identification. Here the parameters tuned iteratively to decorrelate the closed-loop output error.
7. Virtual Reference Feedback Tuning [11] is a fixed structure offline one-shot direct data driven method that can be used to select the controller parameter for the LTI system.
8. Non-iterative data-driven model reference control [2] is a fixed structure method solves the standard identification problem where the input is affected by noise but not the output.
9. Sub-space approaches [2] do not assume any control structure and the system dynamics are represented as a subspace of a finite-dimensional vector space, which consists of the time series data of input/state/output or input/output.
10. Reinforcement learning control [4] models the control system as an agent acting in a Markov decision process and uses Q-learning as the solution of the MDP.

Existing Attempts at Data Driven Control

1. Model Free Adaptive Control (MFAC) [14] was used for unknown discrete-time nonlinear systems with applications.
2. Adaptive dynamic programming (ADP) DDC [15] was utilized for optimal consensus control problem for discrete time multiagent systems with completely unknown dynamics. A data-driven reinforcement learning method is presented using the current and past system data.
3. Event-triggered control scheme [16] was proposed for nonlinear constrained input continuous-time systems based on the optimal policy.
4. Online ADP algorithm [17] is used to learn the optimal solution with partially unknown dynamics, and the identifier network, critic network, and actor network are employed to approximate the unknown drift dynamics, the optimal value, and the optimal policy, respectively.
5. Novel mixed iterative ADP algorithm [18] is proposed to solve the optimal battery energy management and control problem in smart residential microgrid systems.
6. An indirect data-driven trajectory tracking control problem [19] for a class of unknown nonlinear discrete-time systems is presented in item – an approximate model of the controlled object using historical I/O data and neural network, then designs and adjusts the feedback gain matrix online using measured output data and previous estimates.
7. Explicit model predictive control (EMPC) [20] has been demonstrated to be an attractive control strategy in dealing with state constraints and fast dynamics.
8. Model-based method and DDC method can help each other in a complementary manner [21] – Integrated model-data-based zero phase error tracking feedforward control (ZPETFC) strategy is proposed for high-precision motion systems with complex and nonminimum phase (NMP) dynamics.
9. Data-assisted modeling method for motion control of a two-link robotic fish [22] is presented to tackle the unavailability of the complex hydrodynamics thrust mechanism.
10. The heat exchanging unit is widely employed in many industrial processes – difficult to control due to large disturbances. A co-working control scheme [23] is discussed by integrating model-based control method addressing disturbance in the outer-loop supplying water temperature, and a data-driven dual-rate control method aiming to control water temperature and steam flowrate.
11. A lower limb exoskeleton system including mechanical structure and embedded electronic system, and a data-driven repetitive learning control scheme, to address the periodic tracking control issue with learning convergence is presented [24].
12. A data-driven robust output tracking control (DROTC) [25] was used to combine the advantages of sliding mode control and data-driven MFAC for stable pressure control of gas collectors of coke ovens.

Reinforcement Learning as an Approach to Data Driven Control

Reinforcement Learning is highly generalizable – can be used in multitude of problem settings:

1. Deterministic vs Stochastic
 - Some environments will be deterministic while most other real-world environments will be stochastic in nature.
 - Most video games are deterministic in nature, meaning there is direct one-to-one mapping in the I/O.
2. Fully Observable vs Partially Observable
 - Some environments are fully observable and some are partially observable meaning the variable of interest for determining the state is not directly observable and is calculated from some other directly observed variable via techniques such as Kalman filtering or Particle filtering.
3. Time-invariant vs Time-variant
 - Some problem might have environmental parameters which are fixed while in some problems the environmental parameters might change with time either due to some seasonal factors or due to the life cycle of the environment itself.
4. Continuous vs Discrete
 - Both Continuous time as well as discrete event Markov Decision Processes can be solved in a Reinforcement Learning based framework.
5. Single-agent vs Multi-agent
 - Reinforcement Learning framework can be used to model problems with single as well as multiple interacting decision-making entities whose decisions influence the states of one another.
6. Finite-time vs Infinite-time
 - Some control systems might need to be deployed on infinite-time horizon environments and some might need to be deployed on finite-time episodic environments such as games. The time constraint brings in interesting dynamics in the policy the agent learns to optimize its episodic goal.
7. Single-state vs Multi-state
 - Some environments have a single state while most have multiple states which change dynamically according to actions taken by the agent in each state. Multi-state problems are modelled as Markov Decision Processes (MDPs) and solved using Dynamic Programming.
8. Single-objective vs Multi-objective
 - Multi-objective problems can be modelled in RL framework by condensing the multiple objectives into a single objective or giving some sort of goal hierarchy and allowing ample exploration to determine good policies that maximize net cumulative long-term reward.

Sim2Real

Teaching robots in simulation has several advantages:

- Safer for the operational environment
- Avoids breaking or damaging real robots
- Faster than the real-world
- Far less expensive
- Can be run parallelly 'n' number of times

But simulation is only an engineering approximation of reality. So there is a gap between simulation and reality:

- Less accurate in terms of order of representation
- Mostly deterministic – no representation of stochasticities in the system
- Not representative of environmental noise
- Not exhaustive representation of all the operational conditions
- Only as good as our assumptions and understanding of the modelled system

This gap between simulation and reality is a combination of appearance gap and content gap[35]. Appearance gap corresponds to how the environment appears to the learning agent, ie, how realistic the visual data is. And the content gap corresponds to how much well the environment represents the underlying real world, ie, how well the physics is simulated. Appearance gap comes in the simulation of high dimensional input sensors such as camera[36], whereas the content gap comes in the simulation of low dimensional but high-noise sensors such as IMU.

Researchers have proposed several approaches[35] to solve the sim2real problem:

1. System identification
 - Build mathematical model for the physical system – simulator in the context of RL
 - Careful calibration for realistic simulator
 - Calibration is unreliable because many physical parameters of the same machine might vary significantly
 - Due to temperature, humidity, positioning or its wear-and-tear in time
2. Domain adaptation DA
 - Transfer learning to update data distribution in sim to match the real one through mapping or regularization enforced by the task model
 - End-to-end image based RL task DA built on adversarial loss or GAN
3. Domain randomization DR
 - Create a variety of simulated environments with randomized properties and train a model that works across all of them
 - Likely this model can adapt to the real-world environment, as the real system is expected to be one sample in that rich distribution of training variations

Benefits of Data Driven Control

Engineering Benefits

1. New unknown optimal solutions can be discovered using Reinforcement Learning. When given only a singular goal to train against, the agent is given maximum freedom to explore the solution space and thus unimaginable optimal strategies can be discovered.
 - Examples are agents that learn to play games. Atari, Go, Chess and Dota – agents after playing for long enough, through the explore component of their policy can try new strategies in unbiased manner and learn novel strategies from them.
2. Robustness to variations in operating conditions – as long as the data captures the inherent stochasticity in the system and environment.
 - Wide range of scenarios can be simulated digitally or in real experiments and data can be used to learn good behavior in such scenarios. – Take Autonomous Driving for example – it would be impossible to model all the different scenarios a vehicle would be exposed to by hand – only data through experiments and stochastic simulations can come to the rescue.
3. Adaptability – if the time-variation trends are well represented in the data then the control model trained on this data is adaptable to unseen data by extrapolating with appropriate function.
 - Example is a Control System that is deployed in an environment with seasonal characteristics such as Inventory Management System for seasonal goods which will have to adapt to the supply-demand patterns of the good.
4. Theoretical guarantee of capturing arbitrary complexity in the model – neural network used as the function approximator in supervised learning or reinforcement learning setting, give theoretical guarantee of infinite representational capacity which can capture any arbitrary nonlinearity.
 - Control of robotic manipulators – It might be difficult to model the complex high-dimensional kinematics of multi-link robots

Financial Benefits

1. No need to redesign the controller every time it is being designed for a different environmental setting or operating configuration.
 - A data driven control system designed for one set of operating conditions can be further finetuned on data from other set of operating conditions. Example: HVAC control system in Summer and Winter.
2. Data driven control system can better approximate the ground reality and can thus provide more efficient control decision which result in energy and time saving.
 - Example: Traffic light signal control – difficult to model the dynamics by hand – large amount of data collected over all possible scenarios can give much better model
3. A data driven control framework is highly generalizable and knowledge can thus be more easily transferred between two systems.
 - Example: Load scheduling and Flow optimization are similar problems and thus knowledge from one simulation can be used in another simulation and further finetuned for optimality.

Marketing Benefits

1. Data driven models have marketing value associated with them because of the proven utility of ML in more efficient business decision making
2. Data-driven controllers can be pitched as more efficient, robust, adaptive and reliable than conventional model-based controllers

References

- [1] Andrei, N., "Modern Control Theory - A historical perspective", Bucharest, 2005.
- [2] Z. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification, and perspective," *Inf. Sci.*, vol. 235, pp. 3–35, Jun. 2013.
- [3] Z. Hou, H. Gao, F. L. Lewis, "Data-Driven Control and Learning Systems", *IEEE Transactions on Industrial Electronics*, Vol. 64, Iss. 5, May 2017, pages - 4070 – 4075.
- [4] Andrew G. Barto, "Reinforcement learning control", *Current Opinion in Neurobiology*, Volume 4, Issue 6, December 1994, Pages 888-893
- [5] J.C. Spall, Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, *IEEE Transactions on Automatic Control* 37 (3) (1992) 332–341 – SPSA
- [6] Z.S. Hou, The Parameter Identification, Adaptive Control and Model Free Learning Adaptive Control for Nonlinear Systems, PhD dissertation, Northeastern University, Shengyang, China, 1994. – MFAC
- [7] H. Hjalmarsson, Control of nonlinear systems using iterative feedback tuning, in: *Proc. of the 1998 IEEE American Control Conference*, Philadelphia, USA, 1998, pp. 2083–2087. – IFT
- [8] J.G. Ziegler, N.B. Nichols, Optimum settings for automatic controllers, *Transactions of the ASME* 64 (1942) 759–768. – PID
- [9] P.J. Werbos, A menu of designs for reinforcement learning over time, in: W.T. Miller III, R.S. Sutton, P.J. Werbos (Eds.), *Neural Networks for Control*, MIT Press, MA, 1991, pp. 67–95. Approx DP
- [10] P.J. Werbos, Approximate dynamic programming for real-time control and neural modeling, in: D.A. White, D.A. Sofge (Eds.), *Handbook of Intelligent Control*, Van Nostrand Reinhold, New York, 1992. – Approx DP
- [11] G.O. Guardabassi, S.M. Savaresi, Virtual reference direct design method: an off-line approach to data-based control system design, *IEEE Transactions on Automatic Control* 45 (5) (2000) 954–959. – VRFT
- [12] J.X. Xu, Z.S. Hou, On learning control: the state of the art and perspective, *Acta Automatica Sinica* 31 (6) (2005) 943–955. – Iterative Learning Control
- [13] S. Schaal, C.G. Atkeson, Robot juggling: implementation of memory-based learning, *IEEE Control Systems Magazine* 14 (1) (1994) 57–71. – Lazy Learning
- [14] Z. Hou, R. Chi, and H. Gao, "An overview of dynamic linearization based data-driven control and applications," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4076–4090, May 2017.
- [15] H. Zhang, H. Jiang, Y. Luo, and G. Xiao, "Data-driven optimal consensus control for discrete-time multi-agent systems with unknown dynamics using reinforcement learning method," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4091–4100, May 2017.
- [16] Y. Zhu, D. Zhao, and H. He, "Event-triggered optimal control for partially unknown constrained-input systems via adaptive dynamic programming," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4101–4109, May 2017.
- [17] Q. Wei, D. Liu, F. L. Lewis, Y. Liu, and J. Zhang, "Mixed iterative adaptive dynamic programming for optimal battery energy control in smart residential microgrids," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4110–4120, May 2017.
- [18] Z. Wang, R. Lu, F. Gao, and D. Liu, "An indirect datadriven method for trajectory tracking control of a class of nonlinear discrete-time systems," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4121–4129, May 2017.
- [19] A. Chakrabarty, G. T. Buzzard, and S. H. Zak, "Output- tracking quantized explicit nonlinear model predictive control using multiclass support vector machines," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4130–4138, May 2017.
- [20] M. Li, Y. Zhu, K. Yang, C. Hu, and H. Mu, "An integrated model-data-based zero-phase error tracking feedforward control strategy with application to an ultraprecision wafer stage," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4139–4149, May 2017
- [21] S. Verma and J.-X. Xu, "Data-assisted modeling and speed control of a robotic fish," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4150–4157, May 2017.
- [22] Y. Jia and T. Chai, "A data-driven dual-rate control method for a heat exchanging process," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4158–4168, May 2017.
- [23] Y. Yang, L. Ma, and D. Huang, "Development and repetitive learning control of lower limb exoskeleton driven by electrohydraulic actuators," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4169–4178, May 2017.

- [24] J. Qiu, T. Wang, S. Yin, and H. Gao, "Data-based optimal control for networked double-layer industrial processes," IEEE Trans. Ind. Electron., vol. 64, no. 5, pp. 4179–4186, May 2017.
- [25] Y. Weng and X. Gao, "Data-driven robust output tracking control for gas collector pressure system of coke ovens," IEEE Trans. Ind. Electron., vol. 64, no. 5, pp. 4187–4198, May 2017.
- [26] https://en.wikipedia.org/wiki/Classical_control_theory
- [27] https://en.wikipedia.org/wiki/Control_theory
- [28] https://en.wikipedia.org/wiki/Adaptive_control
- [29] https://en.wikipedia.org/wiki/Intelligent_control
- [30] https://en.wikipedia.org/wiki/Optimal_control
- [31] https://en.wikipedia.org/wiki/Robust_control
- [32] https://en.wikipedia.org/wiki/Stochastic_control
- [33] https://en.wikipedia.org/wiki/Data-driven_control_system
- [34] https://en.wikipedia.org/wiki/Machine_learning_control
- [35] <https://lilianweng.github.io/lil-log/2019/05/05/domain-randomization.html>
- [36] <https://medium.com/ai³-theory-practice-business/augmenting-synthetic-images-for-sim2real-policy-transfer-981efd238dd3>

Abbreviations

ADP	Approximate Dynamic Programming
CNN	Convolutional Neural Networks
DDC	Data Driven Control
DNN	Deep Neural Network
DP	Dynamic Programming
DQN	Deep Q Network
IFT	Iterative Feedback Tuning
ILC	Iterative Learning Control
LL	Lazy Learning
MBC	Model Based Control
MDP	Markov Decision Process
ML	Machine Learning
NN	Neural Network
PID	Proportional Integral Derivative
RL	Reinforcement Learning
RNN	Recurrent Neural Network
SPSA	Simultaneous Perturbation Stochastic Approximation
VRFT	Virtual Reference Feedback Tuning

**Siemens
Technolo
gy and
Services
Pvt. Ltd.**
84, Hosur
Road,
Keonics,
Electroni
c City,
Bengalur
u,
Karnatak
a, India
560100

Contact:

1. Saga
r
Path

rudk
ar
(CT
RDA
FOA
ASY-
IN)

2. Saik
at
Muk
herj
ee
(CT
RDA
FOA
ASY-
IN)