LEARNING IN MACHINES

Control of high-tech mechatronic systems traditionally involves feedback and feedforward control, and essentially only uses a few recent measurements. Here, we aim to explore what can be learned from all available sensor data. A general learning framework is developed that exploits the abundance of data of previously executed tasks. Both fundamental insight and experimental results show that such iterative learning control approaches enable substantial performance improvement compared to traditional control. Interestingly, traditional model-based control theory turns out to have an essential role for fast and safe learning from measured data.

TOM OOMEN

Introduction

The learning from data and information has led to impressive achievements in recent years. Computer algorithms are now capable to successfully learn in many domains, including human language, ranging from speech recognition to accurate translations, real-time pattern recognition from images, digital advertising, self-driving vehicles, Atari, and Go [1]. The key enabler has been the availability of large amounts of data as well as ubiquitous and scalable computation and software.

In sharp contrast, high-tech mechatronic systems, such as manufacturing machines and scientific instruments, are often produced and installed with a pre-defined feedforward/feedback control algorithm, and their performance deteriorates over time due to wear, ageing and varying environmental conditions such as temperature variations. Examples range from lithography machines, 2D and 3D printers and pick & place robots, to microscopes and CT scanners.

Interestingly, these high-tech machines are prime examples of mechatronic system design, where control algorithms are typically implemented in a computer environment. Hence, over the lifetime of these high-tech machines, an abundance of data becomes available, yet this is often not exploited to enhance its performance. Indeed, sensors in mechatronic systems are often used for feedback control, which typically only makes use of real-time position and velocity information.

The aim of this article is to explore opportunities for learning from data in machines, possibly from past and already completed tasks, to control them to the limit of their physical capabilities. A framework for fast and safe learning is presented. Furthermore, at the end of the article, several practically relevant questions are addressed, including what should be done for a broad industrial deployment, what performance can be expected for a specific system at hand, and whether learning control can replace traditional feedback controllers.

Learning requirements

Learning in machines imposes several unique requirements, resulting from the fact that such machines are cyber-physical systems, involving interactions with the real world. In particular, the following requirements are considered throughout:

1. Learning should be fast, since machines require experiments in real-time. In addition, fast adaptation can be useful in case of varying operating conditions, e.g., due to temperature changes induced by motor heating or day/night periodicity.
2. Learning should be safe and use operational data, since dedicated experiments may induce production loss and even damage of the machine.

In the forthcoming sections, an approach to learning in machines is investigated that addresses these requirements.

Learning from past tasks

The aim of this section is to investigate the learning from data. This leads to an approach that bridges data-based learning and model-based control.

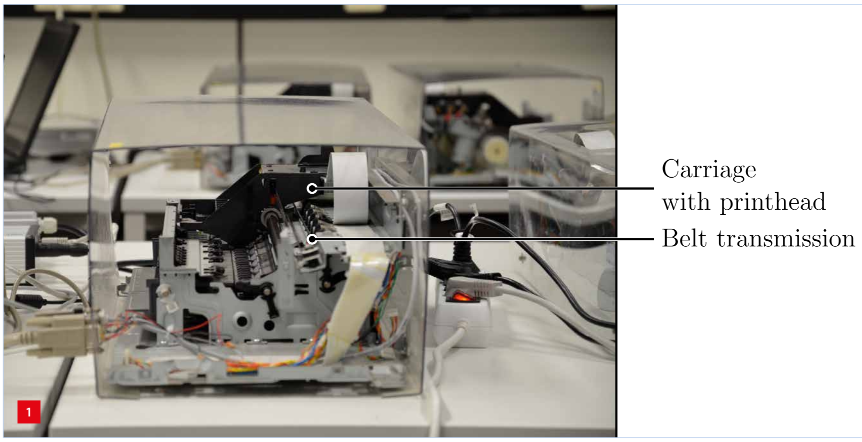
Traditional motion control

The printer in Figure 1 is considered as a key example of a mechatronic system. Here, the goal is to position the carriage that contains the printheads. A motor delivers an input u, which moves the carriage using a belt. The output position of the carriage y is measured using a linear encoder. The printer itself is denoted G.

AUTHOR'S NOTE

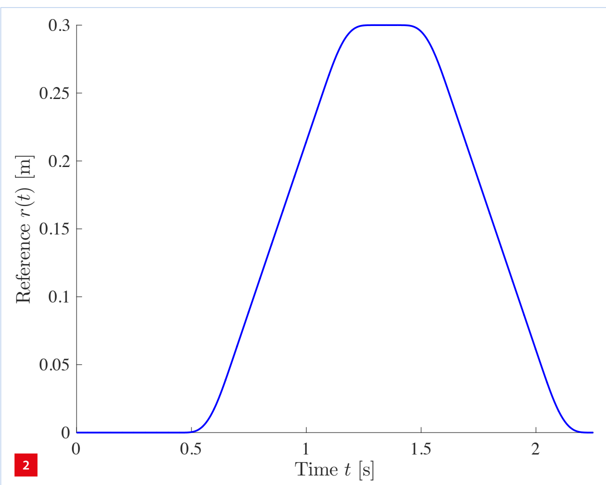
Tom Oomen is associate professor at Eindhoven University of Technology, the Netherlands, in the Control Systems Technology group at the department of Mechanical Engineering.

t.a.e.oomen@tue.nl
www.tue.nl/cst
www.toomen.eu

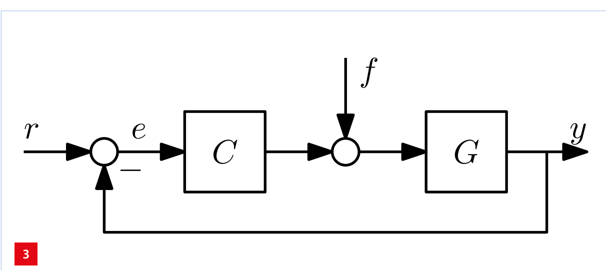


Printer system, used to illustrate traditional motion control and learning.

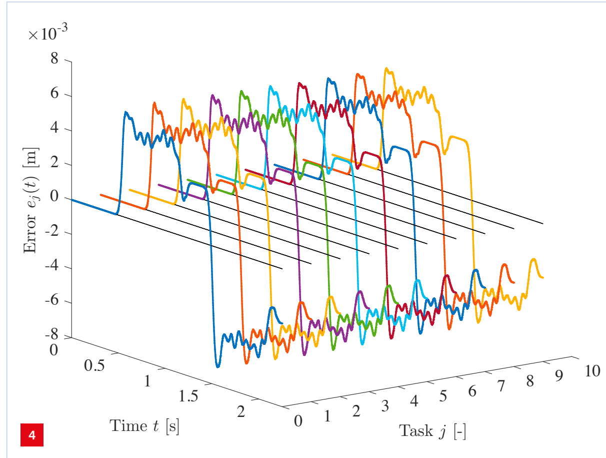
The control task is to track a reference trajectory r, such that the printhead moves over a sheet of paper, see Figure 2. The control problem is thus to choose the control input u such that the error $e = r - y$ is small. Traditionally, this is done using the controller structure shown in Figure 3. Here, C is a feedback controller. Feedforward control is implemented by selecting the signal f. A typical approach is to employ Newton's second law, $f = m{\cdot}a$, where m is an estimate of the mass and $a = d^2r/dt^2$ is the acceleration profile.



Reference for the carriage containing the printheads.



Control architecture, where G denotes the system, C is the feedback controller, f is the feedforward input, r is the reference in Figure 2, and e is the tracking error.



Traditional motion control: the measured error signal is almost identical for subsequent tasks.

Motion control tasks are often performed repetitively. For example, the reference in Figure 2 has to be performed many times before a sheet of paper is printed: during each repetition of the reference in Figure 2, the sheet is moved a few millimeters by a sheet-positioning mechanism. The typical performance of traditional feedback motion control for such repetitive tasks is shown in Figure 4.

Here, ten tasks are shown, where in each task the reference in Figure 2 is tracked. The key observation is that the measured error is almost identical for each task j. Of course, feedforward control by selecting f can lead to a smaller error, but the key observation remains: the error is identical for each task, since the feedforward action f and feedback action Ce do not depend on past errors.

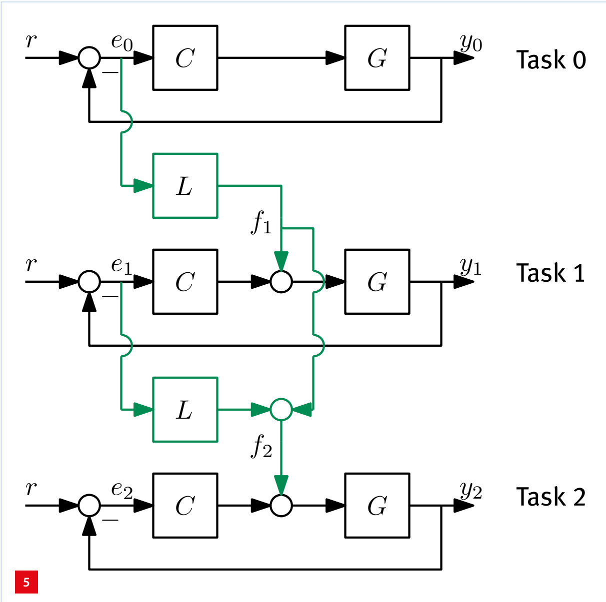
Learning from task to task

The observation that traditional motion controllers lead to a very similar error profile in Figure 4 raises the question: can we learn from past tasks, to improve the performance in the next task, i.e., task $j + 1$? Intuitively, the answer is affirmative: since the error is predictable, it can be compensated for. The practical question is how this can be achieved.

To learn from past tasks, assume that we perform the first task, $j = 0$, with no feedforward, thus $f_0 = 0$. The resulting error during the first task e_0 is then given by $e_0 = Sr$. Here, $S = 1/(1 + GC)$, the so-called sensitivity function, which can be directly derived from Figure 3. Now, consider the following idea. Assume that we measured e_0, but we do not have access to r. What feedforward f_1 should we select to reduce the error e_1? Note from Figure 5 that:

$$e_1 = Sr - GSf_1$$

Next, two key steps are made. First, note that we do not have direct access to Sr, but in fact it was measured in the



Towards learning from data of previous tasks.

earlier experiment: $e_0 = Sr$. Second, let f_1 depend on past errors, for instance:

$$f_1 = Le_0$$

Here, L is a design filter that still has to be specified, see also Figure 5. These steps directly lead to:

$$e_1 = e_0 - GSLe_0$$

This last equation immediately shows that the choice $L = (GS)^{-1}$ leads to $e_1 = 0$.

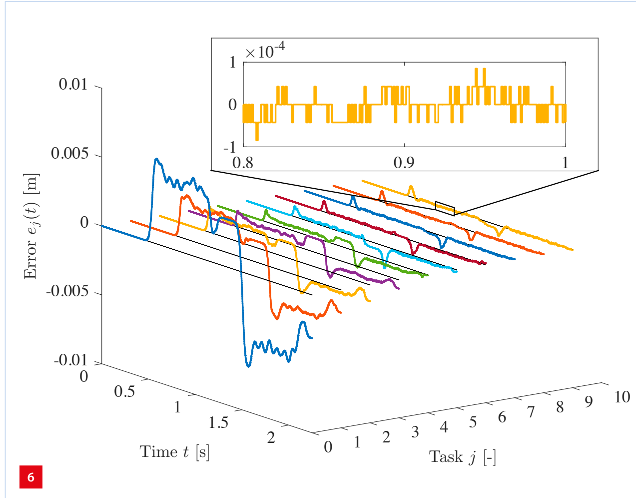
The update law $f_1 = Le_0$ with $L = (GS)^{-1}$ combines the data e_0 with model knowledge GS. Indeed, L is based on a model of the true closed-loop system GS. The key benefit of learning from data is that an approximate model suffices: of course we cannot expect to have access to an exact model of the system. If the model is not exact, then $GSL \neq 1$; so that e_1 is not zero, but typically much smaller than e_0. The central idea is to repeat the learning procedure in the next task $j = 2$:

$$f_2 = f_1 + Le_1$$

This essentially retains f_1 if it is perfect ($e_1 = 0$), and otherwise includes a small correction based on the already small e_1. This is then also done for future tasks:

$$f_{j+1} = f_j + Le_j$$

This idea of updating the control input is referred to as iterative learning control (ILC), see [2] for a historical overview.



Learning from past data in a printer system: fast convergence to encoder resolution.

Experimental results

Application of this procedure to the printer system in Figure 1 leads to the measured error signals in Figure 6. These results reveal impressive control performance: the error is at the level of the encoder resolution after only a few tasks. Hence, this very simple learning update leads to extremely high performance by combining data and model knowledge. Interestingly, these performance levels cannot be achieved using traditional feedforward and feedback controllers due to the presence of significant friction in the system; even though the learning update is a simple linear model it can perfectly compensate for these effects.

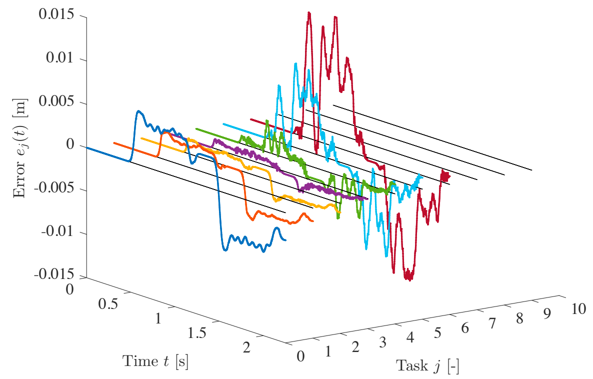
Can learning beat feedback?

Yes! The results in Figure 6 already reveal extremely high performance, which in practice cannot be achieved using traditional feedforward and feedback. The main reason is that feedback is subject to causality. This is well-known, since in $e_0 = Sr$, the term S cannot be made equal to zero due to the Bode Sensitivity Integral, often referred to by control engineers as the waterbed effect. The fundamental reason this integral exists is due to the fact that the physical system G is causal: it only responds to past outputs. In sharp contrast, in learning, one has access to what will happen in the (near) future due to the simple observation that this has been measured in past tasks. In practice, this is done by designing L to be a non-causal filter; practical details are provided in [3].

Fast and safe learning in the face of uncertainty

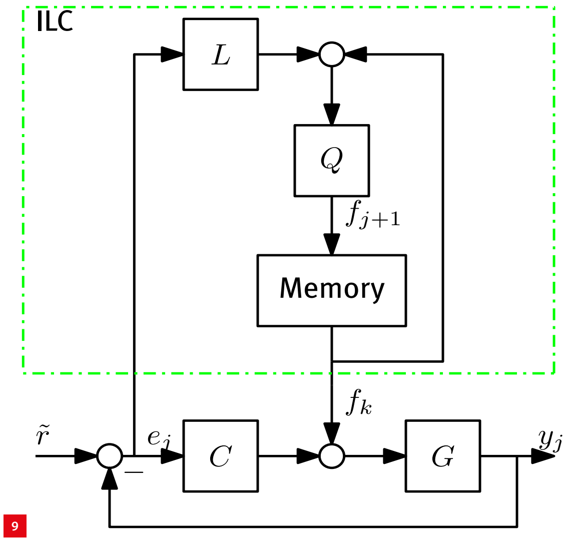
The role of model quality for learning

The results in Figure 6 reveal that the feedforward command signals that result from learning substantially increase control performance. In the previous section, it has been argued that the speed of learning depends on the



7

Learning from past data in a printer system using a simplified printer model (infinite belt stiffness), leading to growing error signals for subsequent tasks. At iteration 7, safety measures necessitated a shutdown of the system.

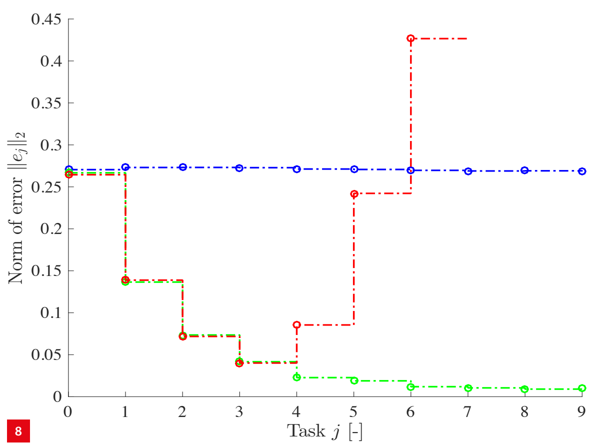


9

Mixed time/task-domain block diagram of iterative learning control, revealing that learning actually is a feedback mechanism.

model quality used to design L. But can it in fact always be guaranteed that the control performance improves?

In Figure 7, the learning procedure of the previous section has been repeated with a slightly different model to determine the learning filter L. In particular, in the experiments of Figure 6, a finite belt stiffness has been assumed, see also Figure 1. In the experiments of Figure 7, a model has been used where the belt is assumed to be infinitely stiff. It is directly observed that in the initial tasks, learning improves performance, but from task 4 onwards, the error actually increases, showing a diverging behaviour until safety measures stop the system at task 7. This can also be seen in Figure 8, where the 2-norm of the error signal is shown for each task, providing a measure of the energy of the error signal. How can feedforward inputs lead to a seemingly unstable system behaviour?



8

Graphs of the 2-norm of the error signal during iteration. Blue: feedback result from Figure 4. Green: learning result from Figure 6. Red: divergent learning behaviour from Figure 7.

Is learning feedforward or feedback?

To understand the behaviour in Figure 7, note that although learning is implemented as feedforward in the time domain, it actually leads to feedback in the task-domain. This can be directly observed in the mixed time/task domain block diagram in Figure 9, where the earlier learning update is obtained if $Q = 1$.

This feedback perspective on learning allows for an explicit analysis of the convergence using control theory. In particular, with the system behaviour $e_j = Sr - GSf_j$ and learning update $f_{j+1} = f_j + Le_j$ it directly follows that:

$$e_{j+1} = (1 - GSL)e_j$$

This type of iteration is ubiquitous in control theory. A very classical result, the Banach fixed-point theorem, implies that this iteration converges in the sense of Figure 8, if the Bode magnitude plot of $(1 - GSL)$ is less than 1 for all frequencies. Thus, convergence, as in Figure 6, 7, and 8, can be directly verified using tools that are traditionally used by mechatronic feedback control engineers. Again, this confirms that learning control in fact is feedback. The feedback perspective on iterative learning from task to task also allows for different choices of L, which can for instance be chosen as a PD controller as in Arimoto ILC approaches [4]. Essentially, this involves a trade-off between required model complexity and convergence speed and behaviour.

Safe learning: the role of robust control

Clearly, when working with physical systems, the diverging behaviour in Figure 7 should be avoided at all cost. This divergence depends on the model that is used for the learning filter L. Control engineers typically have two options in case model errors are too large. First, a better model can be made. Unfortunately, obtaining a model that satisfies the convergence condition for all frequencies requires an extremely high model quality, which is often prohibitively expensive. Second, robustness can be enforced in the design, which is often much more attractive in view of the modelling effort required.

In particular, the field of robust control provides a highly systematic approach for safe learning. Indeed, robustness can be directly enforced by selecting Q in Figure 9. In particular, in case the Bode magnitude plot of $(1 - GSL)$ is not less than 1 for all frequencies, a frequency-dependent Q should be designed such that $Q(1 - GSL)$ is less than 1. Interestingly, this condition can be immediately verified for a set of identified frequency-response functions, which shows a high similarity with traditional mechatronic feedback control design, see [5].

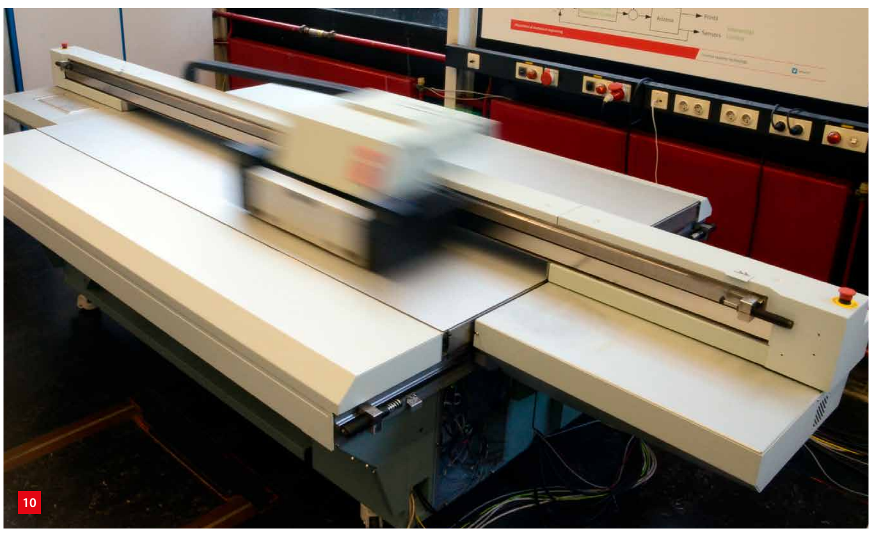
Industrial implementation

The results in Figure 6 reveal an impressive performance improvement. This raises the immediate question: why is learning control not yet standard in industrial mechatronic systems?

Task flexibility

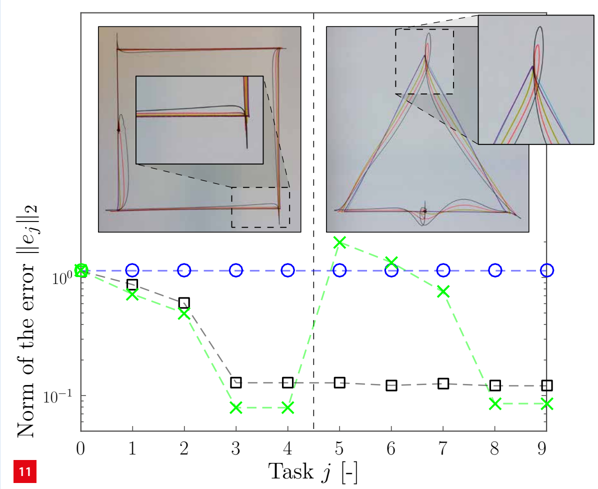
The learning approach outlined in the previous sections assumes that the reference, see Figure 2, is identical for each task. However, in many mechatronic systems the references may change for each task, a typical example being 3D printing. Unfortunately, learning control is highly sensitive for small variations in the reference.

To visualise the troublesome situation, a drawing task has been performed with the 2D industrial flatbed printer in Figure 10. In task 0-4, the goal of the printer is to draw a square. At task 5, the reference is changed to a triangle.



10

Industrial flatbed printer with varying references.



11

Learning with varying references on the 2D flatbed printer in Figure 10. In task 0-4, the goal is to draw a square. From task 5 onwards, the goal is to draw a triangle. Feedback control (blue) leads to mediocre performance. Learning control (green), as described above, leads to an almost perfect square at iteration 3, yet yields very poor performance as soon as the reference changes in task 5. Recently developed algorithms (black) combine task flexibility and high performance through learning.

Clearly, the performance deteriorates significantly, and becomes even worse compared to feedback. Indeed, in case the reference changes each task, it can be shown that feedback outperforms learning.

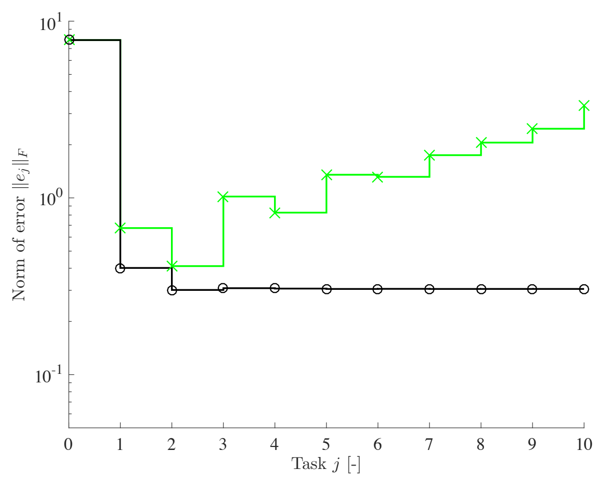
To address these aspects, learning control with flexibility to tasks has recently been investigated, e.g., in [6]. The key idea is to parameterise f_j such that it extrapolates well with changing references. In Figure 11, the potential is already apparent: both flexibility to varying tasks and high performance are achieved with the new approach.

Learning in complex high-tech systems

High-tech systems are becoming increasingly complex. The example system in Figure 1 only has a single input and output, whereas the system in Figure 10 already has three inputs and three outputs. In many high-tech systems, e.g. in lithography, the entire system may have hundreds of inputs and outputs. This raises the question how learning should be performed, and whether the learning approach described above can be applied sequentially or simultaneously for a set of input-output pairs.

Unfortunately, the naive way of learning for a number of input-output pairs often does not work. In Figure 12, it is shown what happens when the learning approach is naively applied to multiple inputs and outputs of the system in Figure 10 simultaneously. Clearly, this may lead to a diverging error, while the individual loops converge.

Interestingly, this aspect directly connects to multivariable control theory. In [7], a unified framework is developed that



12

Iterative learning for the printer system in Figure 10 with multiple inputs. Naively applying the learning approach to multiple inputs and outputs simultaneously (green) leads to divergent behaviour. A systematic approach (black), see the subsection on learning in complex high-tech systems, enforces convergent learning behaviour for complex multivariable systems.

allows a systematic design of multivariable learning controllers for complex systems with many inputs and outputs. Interestingly, the approach focuses on a well-balanced use of models and data. The result in Figure 12 confirms that fast and safe learning is achieved for complex systems.

Data-driven intelligent mechatronic systems

What learning has to offer

Learning enables a major performance improvement in machines by exploiting data from past tasks. A general framework for fast and safe learning has been outlined in this article, enabling intelligent mechatronic systems in the near future that can be controlled to the limits of their reproducible behaviour. The role of model-based approaches has been clearly emphasised to achieve fast learning. Control theory is central to achieve safe learning with convergent error signals, which is an essential aspect for learning in physical systems.

A key remaining question is how much performance improvement can be expected with learning? Also, is a classical feedback controller still required? As a general answer, the field of control is able to compensate to the limit of reproducible behaviour of the physical system under consideration. To investigate what learning has to offer for a particular system, consider the following practical procedure. Perform a sequence of n_{exp} experiments with traditional feedback control and optionally feedforward control implemented and measure the error signals e_j, $j = 0, ..., n_{exp} - 1$, and compute the sample mean, i.e.:

$$m_e = \frac{1}{n_{\text{exp}}} \sum_{j=0}^{n_{\text{exp}}-1} e_j$$

Learning, as has been outlined in the previous sections, is capable of designing a control input that completely compensates for m_e. The performance that can be expected after learning is thus given by signals $e_j - m_e$, $j = 0, ..., n_{exp} - 1$. In this respect, the obtained error at task 10 and beyond in Figure 6 could have been directly predicted from the sample mean of the realisations in Figure 4, where only feedback control is implemented.

The remaining error $e_j - m_e$ is the part that cannot be predicted before the next task starts. Intuitively, feedback control has the task to compensate for these disturbances that occur during the task. Indeed, these disturbances are different each task, but have similar properties for each task, e.g. in terms of their frequency content. It means that as soon as measured data becomes available during the task, a well-tuned feedback controller can effectively address these disturbances. This has been well-known since the advent of optimal control theory in the 1960s: the feedback controller should optimally lead to an error signal which is white noise. In the context of joint learning and feedback, this is investigated in detail in [8]. In conclusion, learning control and a good feedback design are both essential in precision mechatronic systems.

Future developments

In the near future, a further bridge between model-based control and data-based learning is to be expected, which will enable tremendous performance improvements in mechatronic systems. On the one hand, high-tech mechatronic systems are expected to be increasingly complex [9], leading to new learning controllers for multivariable systems [7], unmeasurable performance variables [10], linear parameter-varying dynamics [11], and varying tasks [6]. On the other hand, new developments in control and machine learning will lead to new learning control appoaches, including model-free and reset-free learning [12], kernel-based regression techniques [13], and sparse optimisation [14].

Acknowledgements

This article is the result of a fruitful collaboration with many colleagues in academia and industry, who are gratefully acknowledged. In particular, Frank Boeren, Lennart Blanken, Joost Bolder, Noud Mooren, Robin de Rozario, Nard Strijbosch and Jurgen van Zundert have had a major role in developing and contributing to the ideas outlined here. Financial support is provided by VIDI project number 15698, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO) and ECSEL-2016-1 under grant agreement number 737453 (I-MECH).

REFERENCES

[1] D. Silver, et al., "Mastering the game of Go without human knowledge", *Nature*, vol. 550, pp. 354-359, 2017.

[2] D.A. Bristow, M. Tharayil and A.G. Alleyne, "A Survey of Iterative Learning Control: A Learning-Based Method for High-Performance Tracking Control", *IEEE Control Systems Magazine*, vol. 26, no. 3, pp. 96-114, 2006.

[3] J. van Zundert and T. Oomen, "On Inversion-Based Approaches for Feedforward and ILC", *IFAC Mechatronics*, vol. 50, pp. 282-291, 2018.

[4] S. Arimoto, S. Kawamura and F. Miyazaki, "Bettering Operation of Robots by Learning", *Journal of Robotic Systems*, vol. 1, no. 2, pp. 123-140, 1984.

[5] N. Strijbosch, L. Blanken and T. Oomen, "Frequency Domain Design of Iterative Learning Control and Repetitive Control for Complex Motion Systems", in *IEEJ International Workshop on Sensing, Actuation, Motion Control, and Optimization* (pp. 1-2), Tokyo, Japan, 2018.

[6] L. Blanken, F. Boeren, D. Bruijnen and T. Oomen, "Batch-to-batch rational feedforward control: From iterative learning to identification approaches, with application to a wafer stage", *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 2, pp. 826-837, 2017.

[7] L. Blanken and T. Oomen, "Multivariable iterative learning control design procedures: From decentralized to centralized, illustrated on an industrial printer", *Arxiv: arxiv.org/pdf/1806.08550.pdf*, 2018.

[8] J. van Zundert and T. Oomen, "On Optimal Feedforward and ILC: The Role of Feedback for Optimal Performance and Inferential Control", in *IFAC 2017 World Congress* (pp. 12607-12612), Toulouse, France, 2017.

[9] T. Oomen, "Advanced Motion Control for Precision Mechatronics: Control, Identification, and Learning of Complex Systems", *IEEJ Journal of Industry Applications*, vol. 7, no. 2, pp. 1-14, 2018.

[10] J. Bolder and T. Oomen, "Inferential Iterative Learning Control: A 2D-System Approach", *Automatica*, vol. 71, pp. 247-253, 2016.

[11] R. de Rozario, R. Pelzer, S. Koekebakker and T. Oomen, "Accommodating Trial-Varying Tasks in Iterative Learning Control for LPV Systems, Applied to Printer Sheet Positioning", in *2018 IEEE American Control Conference* (pp. 5213-5218), Milwaukee, Wisconsin, 2018.

[12] R. de Rozario and T. Oomen, "Improving Transient Learning Behaviour in Model-Free Inversion-Based Iterative Control with Application to a Desktop Printer", in *15th International Workshop on Advanced Motion Control* (pp. 455-460), Tokyo, Japan, 2018.

[13] L. Blanken, I. van den Meijdenberg and T. Oomen, "Kernel-Based Regression of Non-Causal Systems for Inverse Model Feedforward Estimation", in *15th International Workshop on Advanced Motion Control* (pp. 461-466), Tokyo, Japan, 2018.

[14] T. Oomen and C.R. Rojas, "Sparse Iterative Learning Control with Application to a Wafer Stage: Achieving Performance, Resource Efficiency, and Task Flexibility", *IFAC Mechatronics*, vol. 47, pp. 134-147, 2017.

Certified training

The content of this article is in part covered by the "Advanced Feedforward Control" training, provided by The High Tech Institute. This course has been certified within the European ECP² framework (see page 60).

WWW.HIGHTECHINSTITUTE.NL/EN/TRAINING/MECHATRONICS