Manufacturing Letters 37 (2023) 45–49

Contents lists available at ScienceDirect

Manufacturing Letters

journal homepage: www.elsevier.com/locate/mfglet

Letters

A human-centered framework to update digital twins

Peter Denno *, Deogratias Kibira

National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899, USA

ARTICLE INFO

Article history:
Received 2 February 2023
Received in revised form 9 June 2023
Accepted 20 June 2023
Available online 13 July 2023

ABSTRACT

Developing a digital twin (DT) involves establishing (1) a predictive capability (a model) relevant to the application, (2) means to collect data from the physical counterpart, and (3) means to apply the collected data to the model. Ideally, with these three goals achieved, long periods of steady-state use of the DT might be interrupted only by failure of the sensors used to collect data from the physical counterpart. In reality, however, it can be difficult to confirm that the DT system occupies this comfortable steady-state position. Assessing uncertainty in the predictive model, and ensuring the relevance of data collected from the physical counterpart are design-time activities with unclear termination points. Distinguishing sensed change in the physical counterpart from sensor failure is a persistent challenge. In this short paper we describe early work towards a human-centered framework to establish, refine, and update digital twins. Condition-based maintenance and gear backlash in production equipment are used as examples.

Published by Elsevier Ltd on behalf of Society of Manufacturing Engineers (SME).

Introduction

This paper describes early work to systematize joint (human/-machine) development of DTs. The work emphasizes the role of human analysts in the creation and maintenance of digital twins. A common use of DT is to support condition-based maintenance (CBM). CBM presents several challenges: (1) CBM may require instrumentation and sensing beyond what is required for ordinary system functions. (2) Inference to an accurate diagnosis can require large amounts of training data [1]. (3) Though CAx models might provide input to a DT, such technology was originally developed for product design and not to represent a product through the changes that occur during its life cycle [2], and (4) simulation technology used is also not well adapted to the repetitive life-cycle of model refinement associated with CBM [3]. For these reasons, it can be costly to use a digital twin in CBM [4]. The work focuses on the key task of determining where in the DT shortcomings reside: in understanding the system's mechanisms (model uncertainty), in data collection (e.g. measurement uncertainty) or heretofore-unrecognized degradation of the physical counterpart. As will be described, we use a Bayesian technique to decide which among these is responsible, and thereby what shortcomings of the DT system to address.

Section 2 of the paper describes the DT refinement cycle. Section 3 discusses the framework. Section 4, the conclusion, provides thoughts on next steps.[1]

The model updating and refinement cycle

Conditions motivating updates to a DT include (a) changes in the physical counterpart, (b) changes in how the analysts understand both causality in the system and shortcomings in modeling that causality, and (c) willingness to invest resources in the effort to improve prognostics and health management (PHM). There are, therefore, dynamic relationships among the DT, its physical counterpart, and the analysts' knowledge. We suggest that there are five forces motivating this dynamic:

- **Observed change in behavior:** These comprise direct observation of the physical system's behavior. Examples include a motor's power consumption for a given task, a battery's energy capacity, its ability to be charged.
- **Observed change in form:** These comprise direct observation of the physical system and its components, for example, a gear's wear, a battery's chemistry.
- **Inferred change in form:** These comprise what might be inferred about distinct system components using established models (here used as submodels) about how those components degrade. For example, a model of how oil in an internal combus-

* Corresponding author.
E-mail addresses: peter.denno@nist.gov (P. Denno), deogratias.kibira@nist.gov (D. Kibira).

[1] The authors are employees of NIST, an agency of the US Government. No external funding sources were used in this work.

https://doi.org/10.1016/j.mfglet.2023.06.002
2213-8463/Published by Elsevier Ltd on behalf of Society of Manufacturing Engineers (SME).

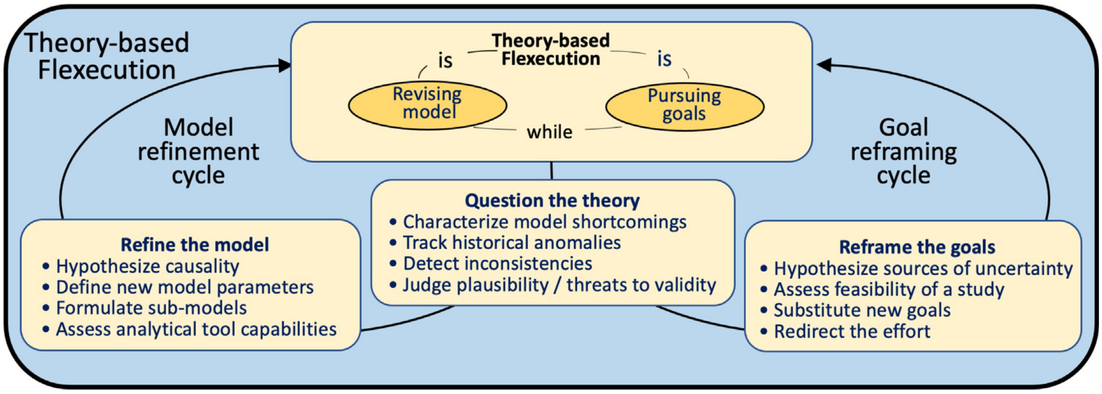

Fig. 1. Theory Flexecution involves the refinement of models in consideration of the theory and opportunities for it improvement. The diagram is adapted from Klein's Flexecution [6] to emphasize the roles of models and theory.

tion engine degrades based on the engine's use can be used to predict the oil's current lubrication properties.

Pressure to explain discrepancy: In as far as the DT's predictive model fails to mimic its physical counterpart, there is a need for its improvement. The changes here include (a) those reflecting better understanding of causal pathways, (b) the accommodation of better sensing of the physical counterpart, and (c) changes in the physical counterpart.

Pressure to optimize: Even when the DT's predictive model is accurately representing the behavior of the physical counterpart and providing actionable recommendations, the computation might be inefficient or the model difficult to update.

There is, therefore, a need to design DTs so that they may be updated easily. Because a system in usage, its DT, and the analysts' understanding are subject to these forces, and because one typically starts with a rudimentary model, it is useful to view DT updating as a dynamic planning problem [5]. Flexecution [6] is dynamic planning by which goals can change based on discoveries made while executing a plan. Klein suggests that in complex settings flexecution is the norm and plans conceived ahead of time and still usable when needed are the exception [6]. Flexecution can be adapted for use in DT refinement. Theory has been described as a hypothesis about the relationship between a collection of models and reality [7]. Theory flexecution, therefore, is flexecution where the goal is to improve a theory. The key activities of theory flexecution are depicted in Fig. 1 below.

The framework

Uncertainties motivate improvement to data collection, model refinement, and model updating. The framework, as depicted in Fig. 2, is a system for assessing uncertainties and opportunities for improvement in the DT system. The key component of the framework is a tool for Bayesian abductive logic programming (BALP) [8]. BALP uses Bayes Nets (BNs) to construct maximum a posteriori (MAP) queries from uncertain observations and statements of uncertain causality, $\mathbf{e}$.

$$MAP(\mathbf{X} \mid \mathbf{e}) = \arg\max_{x} P(\mathbf{X} \mid \mathbf{e}) \tag{1}$$

Each assignment to the random variables $\mathbf{X}$, corresponding to a navigation of the BN, can contribute to a probabilistic explanation of a queried phenomenon. In BALP[2], explanation is hypothetical in the sense that reasoning can be abductive; baseline estimates of likelihood can be used where evidence is lacking. The explanations are mechanistic [10] in the sense that BN inferences can be interpreted as causal chains.

In our work, the *physics-based model of system* in Fig. 2 is a Simulink simulation of the angular velocity of the driven gear on the joint of a robot arm performing a task. The task requires rotation at the joint in order to place the robot end effector at commanded positions.[3] Fig. 3 compares velocity calculated in such simulations of the driven gear under changes of direction and velocity of the driving gear. Two models are compared, that of a healthy gear set and that of a gear set with excessive wear and backlash. When there is a change in the direction of rotation, backlash leads to the driven gear being out of contact with the driving gear. However, because of inertia, the driven gear keeps rotating in the original direction until there is contact again. Then it starts rotating in reverse along with the driving gear. The result is a more jagged graph than the case of zero backlash. For those sections of the curve where there is a sustained rotation in one direction, the plots of velocity are coincident.

The role of the framework is to put explanations such as the above into context with competing explanations. Some of these competing explanations might concern, for example, shortcomings of the model or data collection. As we have suggested in Section 2, motivations for updating concern five forces and the generation of new hypotheses. As Fig. 2 suggests, the key purpose of the BALP analysis (and thereby support for human-centered naturalistic decision making) is to help determine which of three paths (from "Analysis suggests…" in the figure) of refinement to pursue:

1. updating the simulation model, for example either (a) updating the gear wear parameters in the model to better match behavior observed in the physical counterpart, or (b) updating the gear assembly submodel, perhaps replacing the two-gear set with a more realistic model of the physical counterpart's multi-gear gear box [11],
2. improving sensing and data cleaning, for example in the collection of gear velocity, and
3. updating the Bayes net (BN) parameters themselves through supervised learning, this might include, for example (a) distinguishing backlash behavior from behaviors associated with other maladies such as bearing wear or a faulty encoder, and

[2] Similar functionality can be achieved by other means such as hidden Markov models [9].

[3] Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

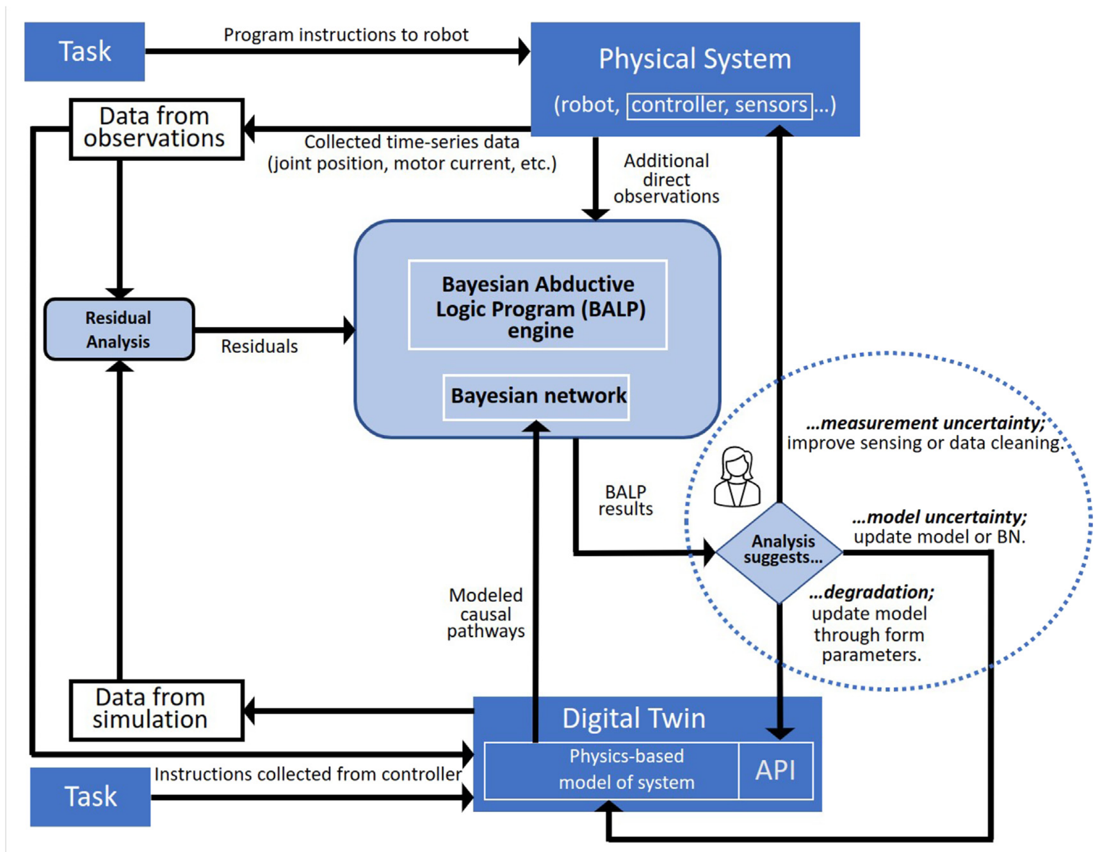

Fig. 2. An architectural view of the framework with annotations showing its use in a robot use case. The role of the human analyst is highlighted in the dotted oval. The analyst is provided probabilities for the three hypotheses and decides whether to improve the model, sensing, or repair the physical counterpart.

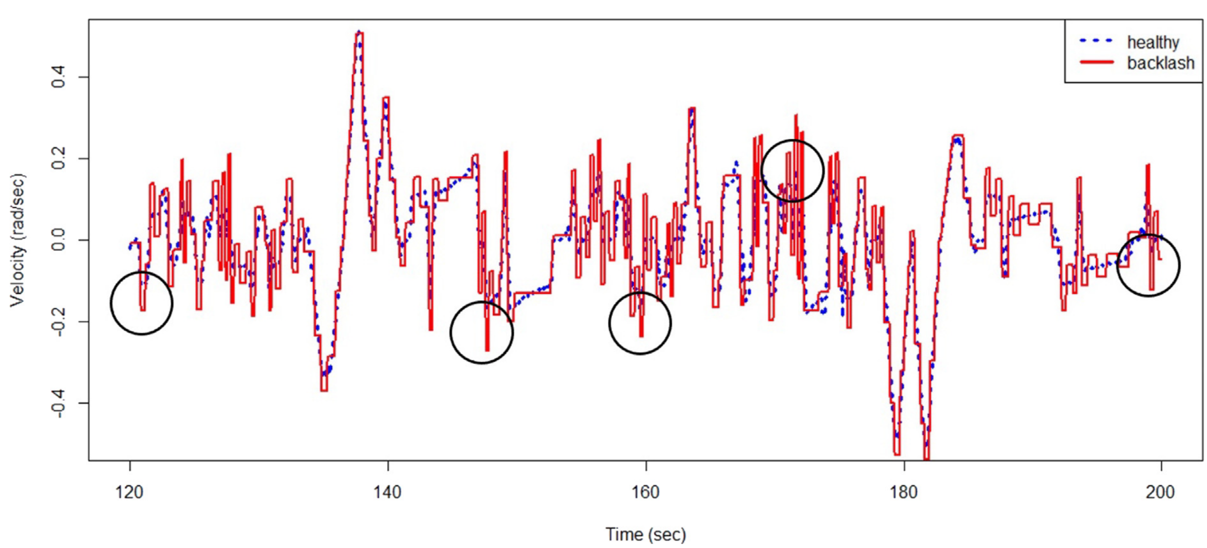

Fig. 3. A comparison of the velocity of driven gear in a joint of the robotic arm under gear backlash versus normal behavior. The circles highlight areas where backlash is most noticeable.

(b) refining the BN model's treatment of the top-level task of attributing discrepancy between model uncertainty, measurement uncertainty and model resemblance to its physical counterpart.

As is typical with BNs, baseline parameterization of the BN relies on expert knowledge. Bayesian rules for two such hypotheses are depicted in Fig. 4. In this example, the goal is to explain the robot's inability to accurately position the tool center point (TCP). The result of BALP analysis depicted at the bottom of the figure depicts two alternative causal chains involving respectively backlash and sensor failure.

The number of possible explanations in such problems can grow exponentially with the number of conditions used in the causal

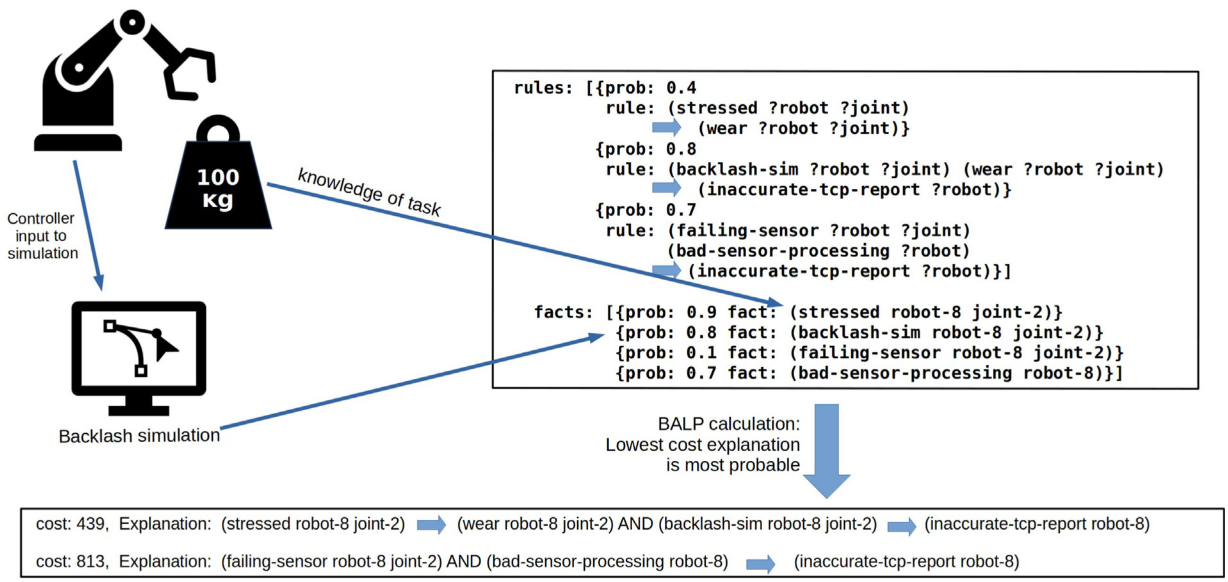

Fig. 4. Two competing explanations for inaccuracy of tool center point positioning: sensor failure versus gear wear and backlash. If confidence in the backlash simulation is sufficient, the likelihood of explanations relying on it are increased.

chains. Experience with the algorithm developed [12] suggests that the MaxSAT solver [13,14] used in the MPE calculation can get bogged down when presented with all hypotheses simultaneously. To keep computation time under a minute, in similar problems we have successfully used tournament selection with games ranking relatively small sets of potential explanations [15].

Our intent is to provide a human-centered experience by appealing to the analysts' desire for a nexus of explanations. Shneiderman [16] has aptly noted that "Human curiosity and desire to understand the world means that humans are devoted to causal explanations, even when there is a complex set of distant and proximate causes for events." The explanations the framework can provide are one step towards our goal. However, a challenge presents itself if the analysts have limited prior experience with analytical tools. Without powerful modern simulation tools, certain hypotheses cannot be easily tested; without confidence in one's abilities with the tool chosen, results should not be relied upon. These concerns can be expressed in the Framework's Bayes net but that alone will not advance the system's ability to do CBM. Another thread of our work focuses on joint (human/AI) formulation of analytical models. In related work [17], we are applying theory flexecution to the joint (human/AI) cognitive work of formulating MiniZinc [18] production scheduling optimizations. It is our intent to use domain-specific languages (DSLs) such as MiniZinc to enable a "conversation" between the analysts and a machine agent supported by large language models (LLMs). For the present work with digital twins we intend to do something similar, replacing MiniZinc with Modelica [19].

Conclusion

We described early work towards a framework for updating digital twins. The key feature of the work is the use of Bayesian Abductive Logic Programming (BALP) to help decide what shortcomings of the DT system to address: model uncertainty, measurement uncertainty, or heretofore-unrecognized degradation of the physical counterpart. Our long-term goal is a framework that supports joint (human/machine) development of digital twins by analysts with limited prior experience. Our work with joint cognitive formulation of scheduling problems suggests that reaching this goal will require detailed collection of data about the analytical model underlying the DT.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Wang Chun Chien, Yaun Kang, Liao Chin Chi. Using Bayesian networks in gear fault diagnosis. Appl Mech Mater 2013;pages 284–287:2416–2420.
[2] Haag Sebastian, Anderl Reiner. Digital twin – proof of concept. Manuf Lett 2018;15(February):64–6.
[3] Kibira Deogratias, Shao Guodong, Weiss Brian A. Building a digital twin for robot workcell prognostics and health management. In: Winter Simulation Conference, Phoenix, USA. p. 12. December.
[4] Brink Harold, Krych Adam, Cardenas Orlando Ramirez, Tiwari Sanchit. Establishing the right analytics-based maintenance strategy. McKinsey & Company; July 2021. p. 6.
[5] Ghallab Malik, Nau Dana S, Traverso Paolo. Automated planning theory and practice. Amsterdam: Elsevier; 2004.
[6] Klein Gary. Flexecution as a paradigm for replanning, Part 1. IEEE Intell Syst 2007;22(5):79–83.
[7] Giere Ronald N. Explaining science: a cognitive approach. Chicago: University of Chicago Press; 1988.
[8] Raghavan Sindhu, Singla Parag, Mooney Raymond J, Raghaven Sindhu, Singla Parag, Mooney J, Raymond. Plan recognition using statistical relational models. In: Sukthanker Gita, Goldman Robert P, Christopher Geib V, Pynadath David, Bui Hung H, editors. Plan, activity, and intent recognition. Amsterdam: Morgan Kaufmann; 2014. p. 57–85.
[9] Koller Daphne, Friedman Nir. Probabilistic graphical models. Cambridge, MA: The MIT Press; 2009.
[10] Glennan Stuart, Illari Phyllis. The Routledge handbook of mechanisms and mechanical philosophy. London: Routledge; 2018.
[11] Aivaliotis P, Kaliakatsos-Georgopoulos D, Makris S. Physics-based modelling of robot's gearbox including non-linear phenomena. Int J Comput Integr Manuf 2023.
[12] Denno Peter. Explainlib: A Clojure library for Bayesian Abductive Logic Programming Identifying the most probable explanation, https://github.com/pdenno/explainlib, 2023.
[13] Ignatiev Alexey, Morgado Antonio, Marques-Silva Joao, Marques-Siilva Joao. RC2: An Efficient MaxSAT Solver. J Satisfiab, Boolean Model Comput 2019;11(1):53–64.
[14] Park James D. Using weighted MAX-SAT engines to solve MPE. In: Proceedings of the National Conference on Artificial Intelligence; 2002. p. 682–7.
[15] Denno Peter. Production scheduling as joint cognitive work. Loughborough University; 2020. PhD thesis.

[16] Shneiderman Ben, Human-Centered AI. Issues in science and technology, 2021. p. 56–61 (Winter).
[17] Denno Peter. Cognitive work in future manufacturing systems: human-centered AI for joint work with models. J Integr Des Process Sci 2023 (in review).
[18] Stuckey Peter J, Marriott Kim, Tack Guido. MiniZinc handbook release 2.2.1. Creative Commons; 2018. Technical report.
[19] Fritzson Peter. Introduction to object-oriented modeling and simulation with modelica using OpenModelica. (c); 2012.