Literature Review: Simulating Flexible Assembly System Event Logs for the Purposes of Process Modelling and Machine Learning – DRAFT

Tero Keski-Valkama

February 28, 2016

1 Topic

Flexible Assembly System is a modular and reconfigurable assembly and tooling workshop with a focus on small and medium-sized batches of varying products. Flexible Assembly System was described formally by Donath and Graves [1] as a system consisting of a set of products each with a specified volume assembled on a workshop consisting of a fixed number of cells.

The assembly steps are performed in cells in parallel. Assemblies and components are transported to the cell where they are combined, tooled or inspected. Products are transported out of the cell for the next assembly step in another cell, or out of the system as end products. The work steps executed in the cells can be manual or automatic. There can be a central storage such as a shelf for storing components, the intermediate assemblies between the steps and the end products waiting to be transported out of the system.

Flexible Assembly System work loads consist of small and medium sized batches where there can be some variation in the products based on customization and personalization. The Flexible Assembly Systems are modular and often composed out of independent modules from different suppliers. The Flexible Assembly System event trace consists of events received from all the separate modules of the system, and additional sensors and triggers added to the system in integration phase or later.

Using a Flexible Assembly System to assemble batches of products generates an electronic event trace which is logged. The events in the event trace are typically at least partially agnostic to the assembly process being performed, and therefore the events do not include tokens connecting the event to a specific final product instance.

Simulating these kinds of event traces provides valuable material for learning systems which can be used to derive assembly process models and to detect deviances. It is in principle possible for an automatic system to detect features in the given material in an unsupervised fashion at least when a human is capable of doing so. A multimodal approach is convenient in data representation when

we need to explore human pattern recognition capabilities. In specific, data visualization as audio allows a human to understand the data melodies, rhythms and possible canon-like features.

2   Description of Method

A proper literature review is a methodological, continuous process. The goal of the literature review is to accumulate a body of relevant existing knowledge about the topic, categorized based on subtopics and keywords. Ultimately the literature review can be presented in the article in a summary form to present the context of the research. The collected references represent a focused area of the existing literature relevant to the object of research.

The review starts from discovery, discovering information sources and starting points of review. The literature review process progresses towards synthesis where the relevant existing knowledge is synthesized together to form an understanding of the composite.

The discovery phase includes a listing of subtopics and keywords, to structure the gathered discovered information into a manageable form. The goal of the discovery phase is to form questions about the existing knowledge and to find answers to them.

The synthesis is composed of the description of the existing knowledge and possible gaps related to the research.

3   Discovery Phase

The topic under review is somewhat cross-discplinary relating to industrial assembly processes, business process modelling and learning systems. Overall the following subtopics are recognized:

1. Assembly process modelling and simulation

   (a) Fault modelling and simulation

2. Mathematical analysis of log data

   (a) Analysing interleaved process traces

   (b) Analysing delays and intervals

3. Process mining

4. Visualization of event logs

3.1   Keywords

The keywords relevant for the research were collected from a set of articles deemed especially relevant for this research. A representative set of articles was read and relevant keywords were picked from titles, abstracts and references. This set of keywords allows for directed browsing of relevant literature.
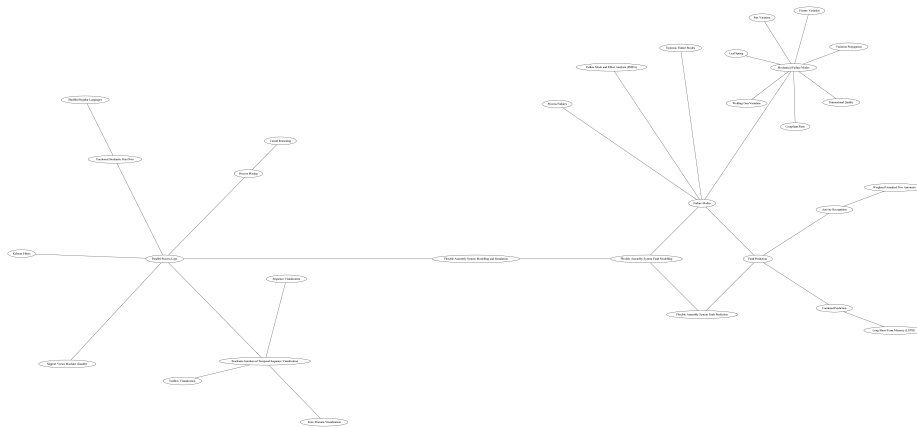
- assembly process

Figure 1: The mind map of the research topic

- assembly system

- complex robotic system

- compliant parts

- dimensional quality

- dimensional variation propagation

- discrete event system

- Failure mode and effects analysis (FMEA)

- failure rates

- failure records

- failure report

- fault detection

- fault mode

- fixture variation

- interleaving

- leaf spring

- multiple failure modes

- Multi-Station Assembly Systems

- outflow visualization

- part variation

- petri nets

- process control

- process failures

- process FMEA

- reduction of irregularities

- reliability

- simulation

- sonification

- time-domain visualization

- uncolored petri nets

- variation propagation

- visualization of sequences

- welding gun variation

The core questions about the existing literature are:

1. How are assembly systems described and modelled?

2. What methods are there to model and simulate assembly processes and related faults?

3. What methods are there to visualize and represent event logs with or without timestamps?

4. What are the relevant keywords and terms to describe this problem space?

4   Synthesis

Assembly systems, construction processes and business processes are often simulated using discrete event simulation, for example in: [2, 3, 4, 5]. These simulations are used for process optimization[6], scheduling and assembly line design. The models of these systems are conveniently described using Process algebra Petri nets (PPNs) [7].

The literature about faults and deviances in assembly processes mainly focus on the faults in the assembled products. There are less academic publications about simulating process deviances in assembly processes such as unexpected process failures, human errors, and deadlocks.

The model of the simulated system is created in production line design [8], by observing the system, or by interviewing the system specialists [9]. The model is changed for each batch of different products being produced.

The fault types relevant in the context of this research are faults in the assembly system itself, not faults in the products being assembled. The assembly system faults can be faults of specific components, or systemic faults, such as deadlocks. Different kinds of faults relevant for the Flexible Assembly Systems have been described for example in [10].

A discrete event simulation output is a sequence of generated events, and the internal system state between the events. The sequence of generated events is useful as a corpus for learning algorithms if it reflects the real conditions and target phenomena well enough. In the case of anomaly detection, the simulation should be representative for the correct operation of the target system, so that deviances are in principle visible in the output data model.

Process mining is the field of inferring the underlying business processes based on observed events and transitions. The process mining methods can be used to compare the supposed process model with the observed process model to detect deviances. Parallel processes create interleaved process traces. The set of all the interleaved process traces form a language. As this language consists of interleaved allowed sequences it is a shuffle language[11].

The methods for describing process traces in the context of process mining are loosely based on the alpha algorithm[12] and as such, they expect process instances to be identified in the activity events to deinterleave the process traces and to infer causal relations of activities. "Event logs need to satisfy two main requirements: (a) events need to be ordered in time and (b) events need to be correlated (i.e., each event needs to refer to a particular case)."[13] There is very little academic literature about process mining of uncorrelated event streams. Uncorrelated event streams are important modelling targets when the event sources are heterogeneous and not necessarily mapped to the specific steps in the known assembly process. This brings event sources such as emergency stop triggers, light ports and smart gateways into the process information system even if the log messages they trigger are not correlated to specific process instances.

Multi-modal representation of the event streams for the purpose of evaluating the human-observable presentation of features in the data is relevant in context of this research. Having a human-detectable pattern in data where the same pattern is challenging to detect with automatic means highlights a potential avenue for additional research.

If a pattern is evident for humans in the data representation, it is in principle discoverable by learning algorithms.

Different human sensory systems are tuned for particular types of patterns, for example music and speech can be appreciated and recognized as audio, but the same signal patterns in visual representation are not recognized as well. A multi-modal representation utilizes different sensory-cognitive capabilities to better facilitate detection of embedded patterns in the data.

There is a lot of literature about discovering structure in sequences and time series by using different visual representations and projections [14, 15]. Visualizations are also used in model-based reporting[16]. Data representation as audio (Sonification [17]) has been investigated for example in [18, 19].

References

[1] M. Donath and R. J. GRAVES, "Flexible assembly systems: an approach for near real-time scheduling and routeing of multiple products," *The International Journal Of Production Research*, vol. 26, no. 12, pp. 1903–1919, 1988.

[2] V. Hlupic and S. Robinson, "Business process modelling and analysis using discrete-event simulation," in *Proceedings of the 30th conference on Winter simulation.* IEEE Computer Society Press, 1998, pp. 1363–1370.

[3] Y. Zhao, C.-B. Yan, Q. Zhao, N. Huang, J. Li, and X. Guan, "Efficient simulation method for general assembly systems with material handling based on aggregated event-scheduling," *Automation Science and Engineering, IEEE Transactions on*, vol. 7, no. 4, pp. 762–775, 2010.

[4] S.-K. Kang, H. Jung, I. H. Im, K.-Y. Chung, and J.-H. Lee, "Active discrete event simulation algorithm using probability distribution of shipbuilding process," in *Information Science and Applications (ICISA), 2013 International Conference on.* IEEE, 2013, pp. 1–3.

[5] A. H. A. Rahnama, "A fuzzy approach on simulating and optimizing the performance of a discrete-event production line," in *Computers and Industrial Engineering (CIE), 2010 40th International Conference on.* IEEE, 2010, pp. 1–5.

[6] N. Sadeghi and A. R. Fayek, "A framework for simulating industrial construction processes," in *Simulation Conference, 2008. WSC 2008. Winter.* IEEE, 2008, pp. 2396–2401.

[7] P. Falkman, B. Lennartson, and K. Andersson, "Specification of production systems using ppn and sequential operation charts," in *Automation Science and Engineering, 2007. CASE 2007. IEEE International Conference on.* IEEE, 2007, pp. 20–25.

[8] H.-J. Bullinger and H. Sauer, "Planning and implementing a flexible assembly system supported by simulation," vol. 25. Taylor & Francis Ltd , Rankine Road, Basingstoke, Hants, RG24 8PR, UK, URL:http://www.tandf.co.uk, 1987, pp. 1625–1634, last updated - 2011-11-11. [Online]. Available: http://search.proquest.com/docview/24612706?accountid=11365

[9] J. A. B. Montevechi and J. D. Friend, "Using a soft systems methodology framework to guide the conceptual modeling process in discrete event simulation," in *Simulation Conference (WSC), Proceedings of the 2012 Winter.* IEEE, 2012, pp. 1–12.

[10] M. Cong, J. Zhang, and W. Qian, "Fault diagnosis system for automated assembly line," in *Intelligent Processing Systems, 1997. ICIPS'97. 1997 IEEE International Conference on*, vol. 2. IEEE, 1997, pp. 1478–1482.

[11] M. Berglund, H. Björklund, and J. Högberg, "Recognizing shuffled languages," in *Language and Automata Theory and Applications.* Springer, 2011, pp. 142–154.

[12] W. Van der Aalst, T. Weijters, and L. Maruster, "Workflow mining: Discovering process models from event logs," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 16, no. 9, pp. 1128–1142, 2004.

[13] W. M. van der Aalst, "Process mining: discovering and improving spaghetti and lasagna processes," in *Computational Intelligence and Data Mining (CIDM), 2011 IEEE Symposium on.* IEEE, 2011.

[14] T. Hein and O. Kramer, "Recognition and visualization of music sequences using self-organizing feature maps," in *KI 2010: Advances in Artificial Intelligence.* Springer, 2010, pp. 160–167.

[15] K. Misue, "Chronoview: A space-efficient method for visualizing temporal patterns," in *Computer Graphics, Imaging and Visualization (CGIV), 2014 11th International Conference on.* IEEE, 2014, pp. 1–4.

[16] M. Schuh, J. Sheppard, S. Strasser, R. Angryk, and C. Izurieta, "An ieee standards-based visualization tool for knowledge discovery in maintenance event sequences," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 28, no. 7, pp. 30–39, 2013.

[17] G. Kramer, *Auditory display*, ser. Santa Fe Institute studies in the sciences of complexity, Proceedings. Reading (Mass.): Addison-Wesley, 1994.

[18] E. S. Yeung, "Pattern recognition by audio representation of multivariate analytical data," *Analytical Chemistry*, vol. 52, no. 7, pp. 1120–1123, 1980.

[19] H. G. Kaper, S. Tipei, and E. Wiebel, "Data sonification and sound visualization," *Computing in science & engineering*, vol. 1, no. 4, pp. 48–58, 1999.