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

Tero Keski-Valkama

February 14, 2016

Keywords. Anomaly detection, Assembly system, Fault detection, Long short-term memory

1 Description of Method

This is a literature review of the first phase of the research, which is about simulating industrial assembly processes and faults for the purposes of process modelling and machine learning research.

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.

2 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 event streams
   (b) Analysing delays and intervals
3. Process mining
4. Visualization of event logs

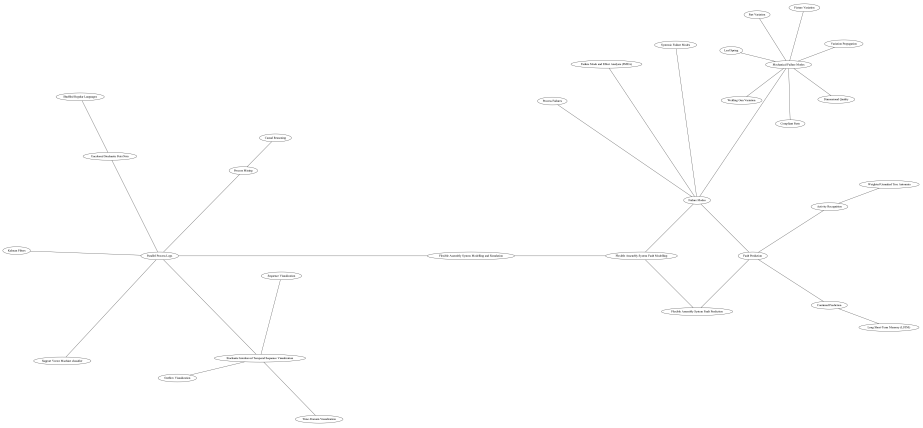

Figure 1: The mind map of the research topic

2.1 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.

- activity recognition
- alpha algorithm
- anomaly detection
- assembly process
- assembly system
- bidirectional LSTM
- causal reasoning
- complex robotic system
- compliant parts
- Connectionist Temporal Classification

- context-free languages
- continual prediction
- Convolutional Neural Network
- corpus
- Deep Recurrent Neural Networks
- dimensional quality
- dimensional variation propagation
- extended hybrid petri nets
- Failure mode and effects analysis (FMEA)
- failure rates
- failure records
- failure report
- fault detection
- fault mode
- fixture variation
- fuzzy generalized stochastic petri nets
- interleaving
- kalman filters
- keyhole plan recognition
- knowledge modelling
- labelling unsegmented sequence data
- leaf spring
- learning Context Sensitive Languages
- Long Short-Term Memory (LSTM)
- multiple failure modes
- Multi-Station Assembly Systems
- nonregular languages
- outflow visualization
- part variation
- petri nets

- plan recognition
- process control
- process failures
- process FMEA
- reduction of irregularities
- reliability
- shuffled languages
- shuffle languages
- shuffle of regular languages
- simulation
- Stochastic Petri Nets
- Support Vector Machine classifier
- system verification
- time-domain visualization
- variation propagation
- visualization of sequences
- Weighted Unranked Tree Automata
- welding gun variation

The core questions about the existing literature are:

1. What are the current best methods of industrial event log based anomaly detection or process mining?
2. What methods are there to model and simulate assembly processes and related faults?
3. What methods are there to visualize event logs with or without timestamps?
4. What methods are there to mathematically model logs generated by parallel processes, or shuffled languages?
5. What are the relevant keywords and terms to describe this problem space?

3 Synthesis

Process mining is the field of inferring the underlying business processes based on observer events and transitions. The current de facto methods for describing process logs in the context of process mining are based on the alpha algorithm and they all expect process instances to be identified in the input event streams.