

Smart manufacturing through a framework for a knowledge-based diagnosis system

Michael P. Brundage¹
Boonserm Kulvantunyou

National Institute of Standards and Technology
Gaithersburg, MD, USA

Toyosi Ademujimi
Badarinath Rakshith

The Pennsylvania State University
State College, PA, USA

ABSTRACT

Various techniques are used to diagnose problems throughout all levels of the organization within the manufacturing industry. Often times, this root cause analysis is ad-hoc with no standard representation for artifacts or terminology (i.e., no standard representation for terms used in techniques such as fishbone diagrams, 5 why's, etc.). Once a problem is diagnosed and alleviated, the results are discarded or stored locally as paper/digital text documents. When the same or similar problem reoccurs with different employees or in a different factory, the whole process has to be repeated without taking advantage of knowledge gained from previous problem(s) and corresponding solution(s). When discussing the diagnosis, personnel may miscommunicate over terms used in the root cause analysis leading to wasted time and errors. This paper presents a framework for a knowledge-based manufacturing diagnosis system that aims to alleviate these miscommunications. By learning from diagnosis methods used in manufacturing and in the medical community, this paper proposes a framework which integrates and formalizes root cause analysis by categorizing faults and failures that span multiple organizational levels. The proposed framework aims to enable manufacturing operations by leveraging machine learning and semantic technologies for the manufacturing system diagnosis. A use case for the manufacture of a bottle opener demonstrates the framework.

INTRODUCTION

Root cause analysis is used in many industries to find the causes of different faults in a system to provide corrective and preventative action (CAPA) plans to alleviate those faults [1-5]. Traditionally, root cause analysis methods in the manufacturing

industry do not present themselves for formal retrieval of information from past studies: it is a one-off practice [1,6-7]. Often, crucial elements of the problem solving process are paper-borne techniques, which do not lend themselves to automated storage and retrieval. The Quality Information Framework (QIF) is beginning to formalize cause and effect retrieval, however it is still free-form text-based, with no formal schema defined [8]. This paper presents a framework for a knowledge-based system for root cause analysis. The knowledge-based method provides a more formal structure for manufacturing system diagnosis. By synthesizing different approaches from engineering and the medical community, this framework allows for more accurate communication, discovery, and reuse of manufacturing diagnosis and corrective and preventative action plans. By providing more structure and learning from the data, manufacturers can reduce the number of misdiagnoses and decrease time to investigate issues.

Currently, the medical industry has a more formal diagnosis procedure than in the manufacturing industry with more research into knowledge-based methods [9-10]. Symptom checker websites, such as 'symcat.com' and 'symptoms.webmd.com,' allow users to input symptoms and output corresponding diseases listed with the probability of occurrence. The goal of this framework is to apply techniques learned from these symptom checker and medical diagnosis systems to provide a manufacturing diagnosis system.

There are four main concepts of medical diagnosis in the medical community [10-11]:

- **Problem assessment** is an evaluation of the patient to assess the current condition,

¹ Corresponding author: mpb1@nist.gov

- **Signs and symptoms** are a sensation or change in health function experienced by a patient,
- **Diseases** are a physiological or psychological dysfunction,
- **Interventions** are an action or series of actions undertaken to respond to the problem.

The process of medical problem assessment includes “Assessing Problem” by “Examining Signs and Symptoms” to “Diagnose Diseases” and “Intervene.” By linking terms from the medical community to the principles of Six Sigma, one can see both the “Assessing Problem” phase and “Examine Signs and Symptoms” are similar to the step “Measure” in Six Sigma, where the performance of the process or system is measured. “Analyze” in Six Sigma is comparable to the “Diagnose Diseases” step because this is the process of discovering the root cause of poor performance and “Improve” in Six Sigma is analogous to the “Intervene” phase since this step is where the root causes are addressed. This paper mainly focuses on the diagnostic aspect (Measure and Analyze) and future work will discuss corrective and preventative action options in length.

While there is no single formal approach for performing root cause analysis throughout all aspects of a factory, there are a variety of techniques currently utilized at different points in the diagnosis process and at various levels of the organization [1,12-16]. For example, root cause analysis is used to diagnose problems of machine failures to help investigate the cause of breakdowns. Alternatively, different techniques can be used at a system level to discover causes of low throughput on a manufacturing line. However, often times once these methods are assessed, this information is not structured in such a way for reuse. The framework presented in this paper addresses that issue.

Six Sigma is a customer-focused, well-defined, problem-solving methodology supported by a handful of structured methods and statistical techniques to reduce variability, defects, and eliminate waste from processes. Six Sigma consists of two main approaches: DMAIC (Define, Measure, Analyze, Improve, Control) and DFSS (Design-For-Six-Sigma) [12]. In DMAIC, during the “Define” phase, the problem, improvement activity, project goals, are defined. “Measure” consists of measuring the process performance, while “Analyze” is determining the root cause of the poor performance. The “Improve” phase consists of addressing and eliminating the root causes and the “Control” phase is where the process is controlled to eliminate errors in the future.

While DMAIC is commonly used for improving existing processes, DFSS is currently used for product development, and is not widely applied to process design and analysis.

A number of case studies targeting improvement of existing processes using Six Sigma have been published. Gijo et al. used the DMAIC approach with the beta correction technique in an automotive part manufacturing company to increase the first pass yield from 94.86 % to 99.48 %, leading to an approximate annual savings of \$87000 [13]. The approach involved a brainstorming session using Ishikawa fishbone diagrams to identify potential root causes followed by data collection and

analysis through a number of statistical analysis methods such as regression analysis, hypothesis testing and Taguchi methods, to validate root causes. A beta correction technique was applied to the process in the control phase for monitoring.

Sharma et al. applied DMAIC to reduce the standard deviation of stub-end hole dimensions during a boring operation in a crankshaft manufacturing process [14]. The dimensional data was collected and analyzed using statistical process control charts (SPC), Ishikawa diagrams, failure modes effect analysis (FMEA) and analysis of variance (ANOVA). Evaluation and hierarchical ordering of potential root causes was based on calculated Risk-Priority-Number (RPN). RPN is a quantitative assessment of risk assigned to a process, usually used in Failure Modes Effect Analysis (FMEA) and is calculated as a product of occurrence, severity and detection [10]. Upon application of a validated solution, the process capability index was improved 56 % and the process performance capability index was improved 353 %.

Kaushik et al. made an attempt to justify the useful role of Six Sigma for small and medium manufacturers (SMEs) by documenting the application of a DMAIC approach to a small scale bicycle chain manufacturer with high rejection rate of cycle chain bushings [15]. Potential root causes were identified with the help of Ishikawa fishbone diagrams and were validated using two-sample t-tests. A design of experiments was conducted in the improvement phase involving validated root causes, to identify optimal values for the process parameters. Incorporation of results into the existing process led to a savings of \$288000 and a decreased Defects-Per-Million-Opportunities (DPMO) level of 0.02. SPC charts were used to monitor the improved process.

Shreeranga et al. employed a DMAIC approach to reduce the number of rejections and rework of automobile leaf springs produced in a small scale foundry [16]. The root causes for rejection and rework were identified through data analysis and the process parameters were optimized to reduce the overall rejection from 48.33 % to 0.79 % leading to savings of \$8000 per year. This was validated through use of gage repeatability and reproducibility (gage R and R) followed by Ishikawa fishbone diagrams to identify root causes and the Taguchi method for optimizing process parameters linked to root causes.

De et al. developed an ontology-based root cause analysis model (RCA) for reducing warranty failure cost [1]. The model utilized information from the Design Failure Mode and Effect analysis (DFMEA), Process Failure Mode and Effect analysis (PFMEA), and corrective action report (CAR) databases to create a Bayesian network which generates all possible failure root causes along with their probabilities of occurrence.

While many studies in root cause analysis for the manufacturing domain exist, currently a formal procedure for capturing this knowledge for reuse does not exist. This paper proposes such a procedure with a knowledge-based framework for manufacturing diagnosis. The case study presented uses

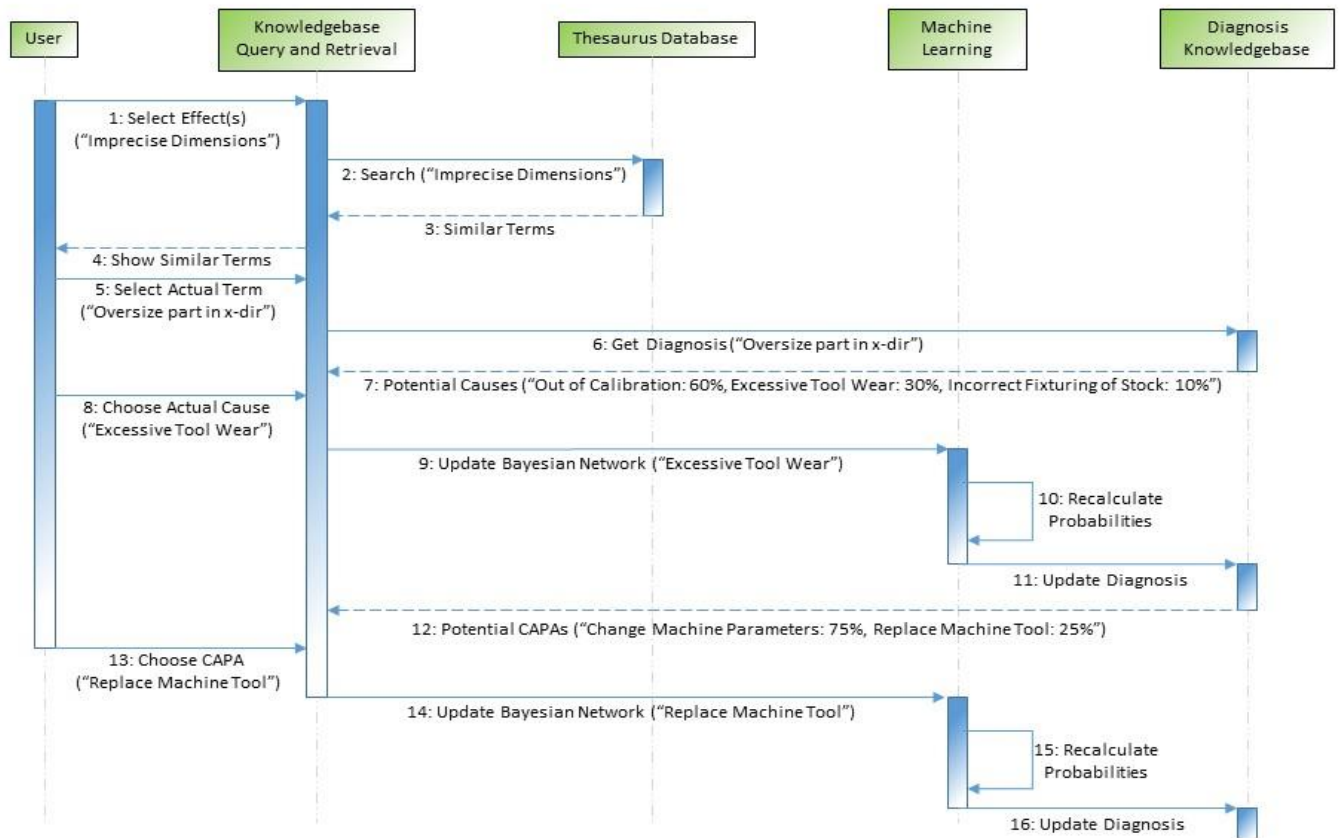


Figure 1: User Interaction with the Manufacturing Knowledge-Based Diagnosis Framework

simulated events data as an input into the system, however the proposed framework is not limited by this constraint. It is possible to use sensor data, maintenance requests, event lists, etc. as input into the diagnosis system.

The rest of the paper is structured as follows. Section 2 is the nomenclature section, while Section 3 presents the proposed framework for the knowledge-based manufacturing diagnosis system. This includes a discussion of the manufacturing diagnosis ontology and the framework architecture. A use case for the manufacture of a bottle opener is studied in Section 4 to demonstrate the framework. Lastly, conclusions and future work are discussed in Section 5.

NOMENCLATURE

The nomenclature section is broken up into two subsections, one for the general text of the paper, and one specifically for the use case example.

General Nomenclature

ANOVA	Analysis of Variance
CAR	Corrective Action Report
CAPA	Corrective And Preventative Action
DFSS	Design for Six Sigma
DFMEA	Design Failure Mode and Effect Analysis

DMAIC	Define-Measure-Analyze-Improve-Control
FMEA	Failure Modes Effect Analysis
PFMEA	Process Failure Mode and Effect Analysis
QIF	Quality Information Framework
RCA	Root Cause Analysis
SME	Small and Medium Manufacturer
SPC	Statistical Process Control

Use Case Nomenclature

For example, LTH FM_B stands for Low Throughput for Front Machining of Body.

FM	Front of Bottle Opener Machining Process
LTH	Low Throughput
HDT	High Downtime
HOOS	High Out-of-Specification Parts
HSR	High Scrap Rate
HRR	High Rework Rate
ON	Operator Negligence
WJC	Water Jet Cutting Process
WMAT	Wrong Material Property
WMS	Wrong Machine Setting
_B	Machining of Body of Bottle Opener

MANUFACTURING DIAGNOSIS FRAMEWORK

This section presents the system analysis artifacts that populate the proposed knowledge-based diagnosis system. First, a sequence diagram in Figure 1 illustrates an example of how a user may interact with the diagnosis system.

In the figure, a user inputs effect(s) (e.g. symptoms) that are observed, which are compared against a thesaurus of previously used terms. Once these effect(s) are selected, potential causes (e.g. diseases) are retrieved from the diagnosis knowledge base with probabilities of occurrence. When a root cause is selected, the probability of the diagnosis in the knowledge base is updated using machine learning. The system presents and updates probabilities of a related corrective and preventative action associated with the diagnosis in the same way. Several functional components enable the diagnosis system that addresses the requirements outlined in the introduction section. One key component is the ontology, which allows the diagnosis knowledge to be represented. The next subsection describes the core concepts of the ontology.

Manufacturing Diagnosis Ontology

The purpose of the concepts in the manufacturing diagnosis ontology is to enable a declarative representation of Failure Mode and Effect Analysis (FMEA), Root Cause Analysis, and Symptom-Diagnosis-Treatment types of information. Figure 2 illustrates five key abstractions of the ontology: *Problem*, *Diagnosis*, *Corrective and Preventative Action (CAPA)*, *Effect*, and *Cause*. In manufacturing, often times failure mode and effect (or cause and effect) are chains of the same concept and consequently the proposed manufacturing diagnosis ontology abstracts them as *Problem*. In a diagnosis-knowledge assertion, the *Problem* can play a role of a *Cause* (failure mode) or *Effect* (the failure). For example, “Out of Specification Part”, “Excessive Tool Wear”, and “Wrong Process Parameter” can all be viewed as the *Problem*, but the “Excessive Tool Wear” can play both *Cause* and *Effect* roles in the following diagnosis relations, *Diagnosis*(*Effect*: “Out of Specification Part”, *Cause*: “Excessive Tool Wear”), *Diagnosis*(*Effect*: “Excessive Tool Wear”, *Cause*: “Wrong Process Parameters”).

Figure 2 shows the relationships between these core concepts using a Unified Modeling Language (UML) class diagram [17]. The *Diagnosis* concept is a piece of knowledge that characterizes the *Cause* and the *Effect* relationships (and therefore they are association classes). An instance of *Diagnosis* can link between multiple occurrences of *Cause* and *Effect*; however, they will be interpreted to have conjunctive relationships. In other words, an assertion of *Diagnosis* knowledge with multiple occurrences of *Cause* and multiple occurrences of *Effect* means that those causes altogether result in those effects altogether. For example, the intention of an assertion, *Diagnosis* (*Effect* (“Low Throughput”), *Cause* (“Frequent Breakdown”, “High Mean Time To Repair”)), is to say that if both the “Frequent Breakdown” and “High Mean Time To Repair” occur together, the result is “Low Throughput,” and implies that both should be addressed together.

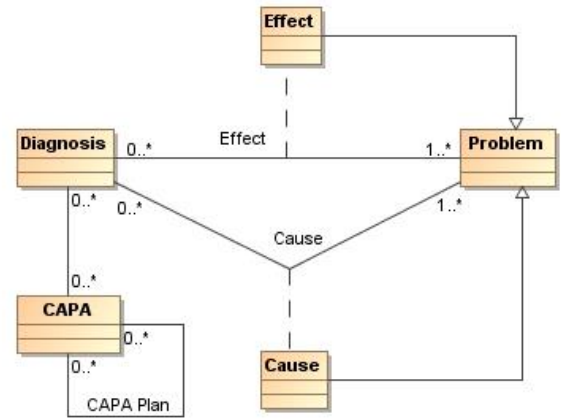


Figure 2: Manufacturing Diagnosis Ontology

Each *Diagnosis* also has a relationship with *CAPA*, which is a method or solution to remedy the situation represented by the *Diagnosis*. In traditional FMEA analysis, *CAPA* may be viewed as associated with *Cause*. In the manufacturing diagnosis ontology, *CAPA* has a direct relationship with *Diagnosis*. Each *Diagnosis* assertion may have one or more associated occurrences of *CAPA* that are also interpreted as conjunctive relationships. In other words, all those specified corrective and preventative actions are necessary to capture the diagnosis and that an alternative corrective and preventative action of the same cause and effect would be captured in another instance of *Diagnosis*.

A corrective and preventative action may be reused in multiple diagnoses. They can also be aggregated to create a higher level corrective and preventative action, called a *CAPA Plan*. The ontology intentionally does not provide a construct to capture a sequence of corrective and preventative actions because the knowledge base does not intend to capture an executable process. If it becomes necessary to capture a sequence of corrective and preventative actions, the higher level corrective and preventative action plan can specify a description of the necessary sequence. The chain of the cause-effect relationships is however captured implicitly through the diagnosis assertions.

Other functional components support the instantiation and usage of the ontology. In the next subsection, the overall functional architecture is presented outlining other functional components, and how they work with the ontology and interact with a user.

Design of a Functional Architecture

Figure 3 illustrates the functional components in the framework. The dotted lines represent the information flows between components. The solid lines represent a functional flow or a trigger, i.e., the component at the beginning utilizes the component at the end of the arrow. The functional flow may also include an information flow. The Thesaurus of Symptoms database stores alternative terms (i.e., labels) for symptoms. This addresses the issue of different terms being used in different organizations or factories for the same items and improves

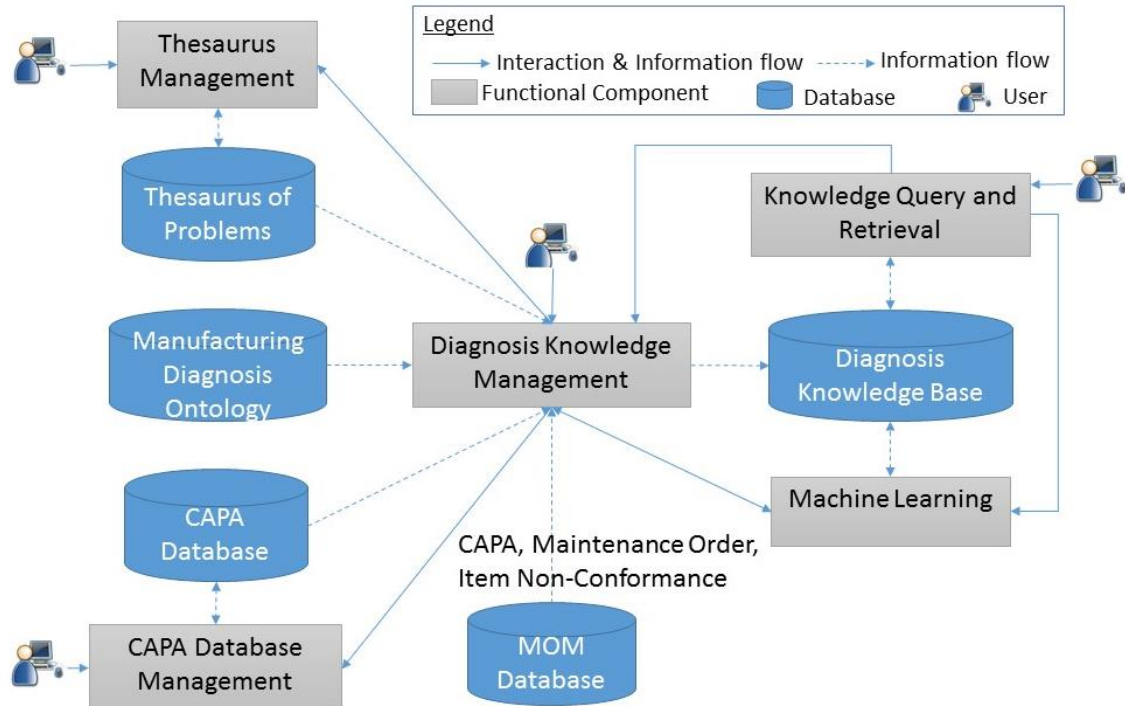


Figure 3: Functional Architecture

knowledge retrieval. The Thesaurus is expected to be evolving. A user drives its creation and maintenance through the Thesaurus Management component. It may provide additional functions such as corpus analysis and a governance process to help a user extract terms and harmonize them across organizations.

The Diagnosis Ontology database stores the concepts described in the previous subsection. Additional concepts may be needed to help formally describe the diagnosis and organize/classify the diagnoses. For example, for the same cause and effect different corrective and preventative actions may be needed depending on the environment (e.g., type of equipment, high humidity or low humidity day, to what kind of product it applies.) Formally capturing environmental conditions may help a user better identify a specific corrective and preventative action. Concepts such as Ishikawa 6M [11-12], [14-15], and ISA-95 functional hierarchy [18] or the ISA-88 [19] physical model may be used to classify the diagnoses. This hierarchy consists of five levels: enterprise, site, area, work center, and work unit. The enterprise level determines what products will be manufactured, at which manufacturing site, and how much will be manufactured. The site level is the physical location determined by the enterprise. It can consist of areas, production lines, production cells, and production units. An area contains the work centers, which consist of process cells, production units, production lines, and storage zones. Lastly, the work centers and work units are the equipment used to manufacture the product.

The CAPA Database stores details about a corrective and preventative action as described in the ontology and may include information beyond the associated method such as use case examples, related documents, and software or hardware tools. A

user drives the creation and maintenance of such information through the CAPA Database Management component.

The Thesaurus of Symptoms, Diagnosis Ontology, and CAPA Database are utilized by a user of the Diagnosis Knowledge Management module to create the diagnosis knowledge, which is maintained in the Diagnosis Knowledge Base. A user may create diagnosis knowledge directly based on his/her prior knowledge. On the other hand, the management component can provide functionality to help glean the diagnosis knowledge from a Manufacturing Operation Management (MOM) Database [18]. The MOM level of a manufacturing facility includes the site, area, work center, and work unit as listed in the hierarchy earlier. This level contains resources such as personnel, equipment, and materials, as well as other documents that are required for completing work. Such information may be available through transaction types, such as Corrective and Preventive Actions, Maintenance Orders, or Items Non-Conformance. When a user creates a new piece of diagnosis knowledge, he/she may have a need to add or change the details of a symptom or a corrective and preventative action and trigger the functions in the Thesaurus Management or the CAPA Database Management component, respectively.

The Machine Learning module interacts with the Diagnosis Knowledge Management component to update the probabilities of occurrence within the network of diagnosis knowledge.

The end user of the framework interacts with the Knowledge Query and Retrieval module to search and enter problem(s) he/she has encountered and navigate through the network of relevant diagnosis knowledge to find the probable causes based on a probability learned by the Machine Learning module. If a

user agrees with a particular diagnosis, it may trigger the Machine Learning module to update the probabilities within the network as well. If a user believes that a new diagnosis needs to be added or an existing one needs to be updated, he/she can invoke the Diagnosis Knowledge Management component to add or change the diagnosis knowledge. The Knowledge Query and Retrieval module helps a user navigate from the cause to the corrective and preventative action in a similar way.

In the next section, we present a use case that illustrates how the framework may be used for manufacturing diagnosis.

USE CASE EXAMPLE

To demonstrate the diagnosis framework, a use case involving the manufacturing of a bottle opener, the “Lion’s Jaw,” was developed. The Lion’s Jaw bottle opener is illustrated in Figure 4.

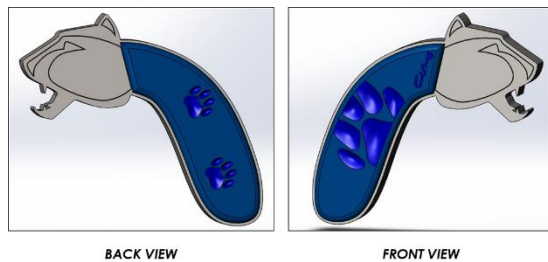


Figure 4: Lion’s Jaw Bottle Opener

The bottle opener was produced in the Factory for Advanced Manufacturing Education (FAME) lab, at the department of Industrial and Manufacturing Engineering, The Pennsylvania State University. In this effort, the manufacturing diagnosis framework was leveraged in concert with a high-fidelity simulation model for identification of common root causes related to the production plan and quality deviations encountered during manufacturing of the Lion’s Jaw product. The Lion’s Jaw manufacturing flow consists of a number of machining and secondary finishing processes. The manufacturing flow of the

Lion’s Jaw bottle opener is based on the job shop process model and has a number of diverse processes, typically found in an SME.

Manufacturing Flow for the Lion’s Jaw

The Lion’s Jaw is an assembly of 2 distinct components: Body and Handle, each having its own manufacturing cycle. Figure 5 illustrates the overall manufacturing flow cycle for the body and handle. The manufacturing cycle for both the aluminum handle and steel body begin with the blanks produced by water jet cutting from a plate. The process for producing the body starts from machining of the front side and front pocket features on the blanks. Downstream, machining of the back side is performed to create another pocket. Further downstream, the machined bodies are bead blasted to produce a matte finish. Four of the bead blasted bodies are then held in a fixture for Wire EDM. In this process, the crucial geometry required for opening the bottles is produced. Finally, two bodies are held in a fixture for engraving the Lion’s face.

The manufacturing cycle for the aluminum handles is similar to that of the body. Two aluminum handle blanks produced using water jet cutting are held in the special photo-adhesive fixture for front side machining. Next, the two semi-processed handles are placed in a soft jaw fixture for back side machining. Machined handles are then anodized to impart specific color. (Blue is commonly used). Finally, the anodized handles are then held in a soft jaw fixture for engraving.

The last stage of this process involves assembling a body with two handles using specialized glue to hold the handles in place.

Simulation Model of Lion’s Jaw Manufacturing Line

A simulation model for the manufacturing of the Lion’s Jaw Bottle Opener was built in Simio, a simulation, production planning and scheduling software. Each process was modelled as a work station with process parameters such as processing

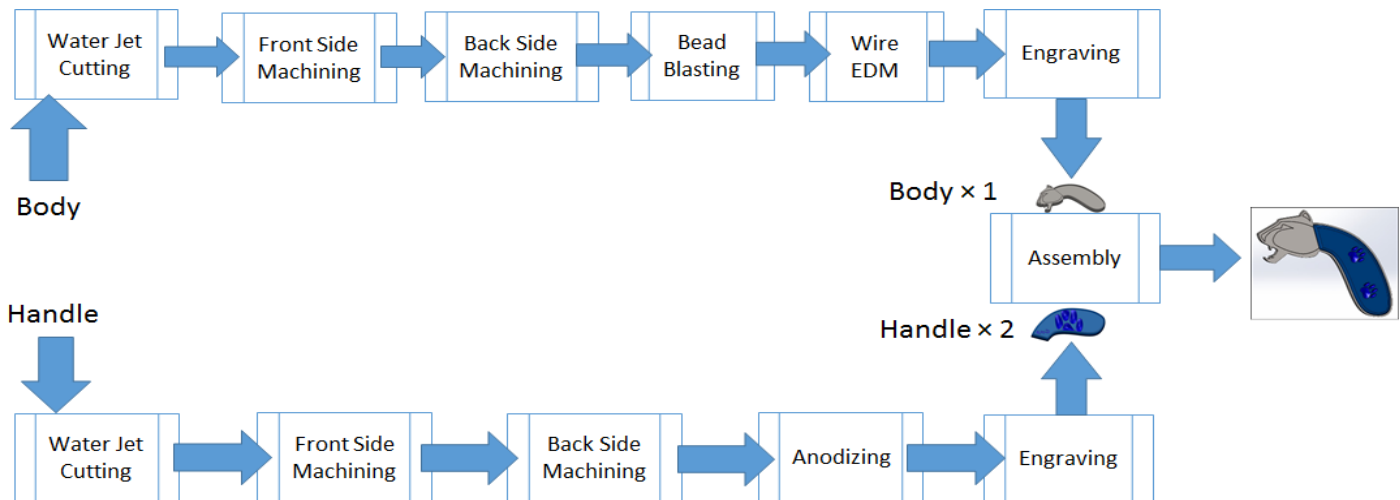


Figure 5: Lion’s Jaw Bottle Opener: Manufacturing Process Flow

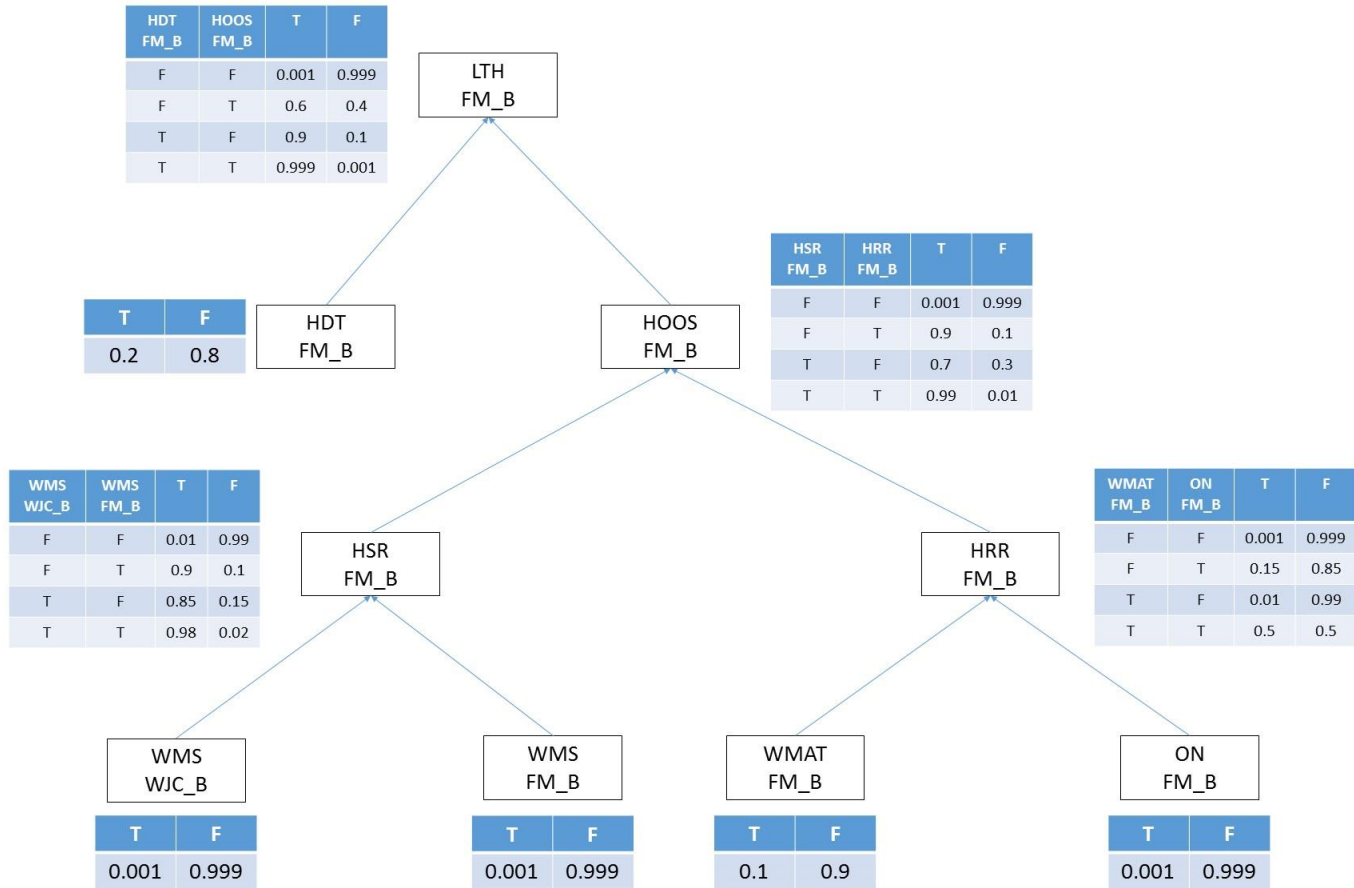


Figure 6: Bayesian Network for Lion's Jaw Bottle Opener Manufacturing

times, down times, batch sizes, etc. obtained using actual cycle time studies and process planning data, in an effort to mimic the real world production process. ISO 22400 key performance indicators (KPIs) are used for monitoring the operations at level 3 of the ISA-95 framework, i.e., manufacturing execution systems level [20-21]. The simulation model was built to compute and produce some of the KPIs such as rework ratio, scrap ratio, throughput, utilization, availability, etc. as per the ISO 22400 standard for each sub-process in the Lion's Jaw manufacturing cycle. These KPIs could be then used to identify certain problems in the manufacturing cycle such as high rework, high throughput, high WIP etc. The frequencies of occurrences (probabilities) of different problems or causes, including the system and process level causes, were recorded along with their observed effects. For example, during the simulation run, whenever the throughput dropped below the specified threshold for low throughput, a "low throughput" effect is said to have occurred.

The possible causes were determined through discussions with the machine operators and observations of the system. The states of the possible causes (i.e., whether the causes were occurring (state=true) or not (state=false)) were simulated to trigger the low throughput event as well as the number of times each cause was responsible for the effect. A binary approach was used to classify all continuous variables (causes and effect) to

true/false by discretization (e.g., an effect of high out-of-specification parts (HOOS) occurred when the ratio of out-of-specification parts to good parts exceeded a threshold of 0.4 for all processes). However, the system is not limited to binary states. The problems are also classified according to the ISA-95 equipment level where they occurred.

Some of the problems that were modeled in the simulation are low throughput, high cycle time and high work-in-process (WIP) at the overall manufacturing system (area level); and low machine throughput, scrap rate, out-of-specification parts rate, rework rate, operator negligence, high machine downtime, and wrong machine settings (equipment level) at each production line and at each machine.

Training the Bayesian Network

The Bayesian representation of part of the Lion's jaw manufacturing process (Front Machining process of Body) is shown in Figure 6 below for the low throughput (LTH FM_B) effect and its possible causes, e.g. high out-of-specification parts, as well as possible sub-causes like high rework rate and high scrap rate. The inputs to the network are probability tables for each node as seen in Figure 6. These probabilities were obtained from the Simio simulation model. A Bayesian network was selected because it models the cause-effect (causal) relationship between nodes which increases the accuracy of the system.

Future work will investigate other machine learning techniques and their applicability to the manufacturing diagnosis framework. The output of the Bayesian network is the probability of a root cause if a particular effect was observed. For example, the probability of WMS FM_B being the root cause given that we observed effect LTH FM_B, or ON FM_B being the root cause given that we observed effect LTH FM_B.

Demonstrating the Diagnosis Framework

To demonstrate the framework, a user inputs the effects (symptoms) he/she is observing. The system will suggest similar effects based on what is stored in the thesaurus. In this case the user inputs “Low Throughput – Front Machining Body.” Once an effect is selected from the thesaurus, the system will present possible root causes and the probability of occurrence. The results from the Bayesian network in Figure 6 are presented in Table 1.

Table 1: Use Case Results

	WMS WJC_B	WMS FM_B	WMAT FM_B	ON FM_B	HDT FM_B
LTH FM_B	0	0	0.33	0	0.67
HRR FM_B	-	-	0.85	0.15	-
HSR FM_B	0.47	0.53	-	-	-
HOOS FM_B	0.16	0.2	0.57	0.07	-

If the user inputs low throughput (LTH FM_B, the framework outputs that it is caused by high down time (HDT FM_B) of the machine 67% of the time and wrong material property (WMAT FM_B) 33% of the time. If high rework rate (HRR FM_B) is observed, it is due to wrong material property (WMAT FM_B) with a probability of 0.85 and operator negligence (ON FM_B) with a probability of 0.15.

Once a user investigates the cause(s) and properly diagnoses the problem, the system will update the probabilities for future diagnosis. A user will then be provided with a list of possible corrective and preventative actions, which will be illustrated in future work.

By using this framework, a user can save time and effort in investigating problems on the manufacturing floor. In this case study, when the manufacturer observes a problem (for example “low throughput”), he/she can very quickly observe that this is most often caused by high down time. This can allow the manufacturer to first investigate this issue since it is most likely to be the cause of the problem. Once the root cause of the problem is determined, the probabilities are updated. Future work will also incorporate other decision variables, such as the time to investigate the problem and the cost to alleviate the issue. This gives the manufacturer more flexibility if they want to first investigate the problem with the lowest cost or the problem that takes the least amount of time to diagnose. Multi-criteria decision making techniques will be investigated to aid in this process.

CONCLUSIONS AND FUTURE WORK

This paper introduces a framework for formal, systematic manufacturing diagnosis of problems arising in manufacturing systems. A functional architecture is presented that consists of many different components. The manufacturing diagnosis

ontology enables a declarative representation of root cause analysis to capture cause and effect relationships. A thesaurus of common terms is used to alleviate the miscommunication problem across organizations. The machine learning module provides predictions of diagnoses based on previous experiences. Lastly, the knowledge base stores this knowledge for reuse. A use case example is conducted to demonstrate the framework with the manufacturing of a bottle opener.

Currently, the data for the system is collected through observations of experts (in this case machine operators), however in the future this research will study how to solicit input from both users and existing databases of maintenance issues. While the use case presented here is for one manufacturing line, it is important to understand how this system will scale to larger lines, plants, and multiple facilities. Future work will study the use of crowdsourcing to expand the knowledge for sharing privately among different facilities within the same enterprise or publicly across enterprises. It is necessary to analyze what information is necessary to allow for accurate diagnosis among different facilities. The ontology will be expanded to accommodate the requirement for multiple factories.

Another area of study is the use of this tool for prognosis. This tool is presented for diagnostic purposes (to find the root cause once a problem is already occurring), however future work will study the applicability to better predict problems before they occur. On top of this, the use of this tool in real time (e.g. using sensor data to trigger the manufacturing diagnosis framework instead of user input) will be studied to allow for quicker diagnosis. The machine learning module will be updated to deal with on-line data.

DISCLAIMER

The use of any products in this paper does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose.

ACKNOWLEDGMENTS

The authors would like to thank Vittal Prabhu, The Pennsylvania State University, and Farhad Ameri, Texas State University, for their assistance in writing the paper. The authors would also like to thank Michael Sharp, NIST, for his technical advice with machine learning techniques.

REFERENCES

- [1] De, S., Das, A., & Sureka, A. Product failure root cause analysis during warranty analysis for integrated product design and quality improvement for early results in downturn economy. *International Journal of Product Development*, 2010. 12(3-4), 235-253.
- [2] Mahto, D., & Kumar, A. (2008). Application of root cause analysis in improvement of product quality and productivity. *Journal of Industrial Engineering and Management*, 1(2), 16-53.
- [3] Dorsch, J. J., Yasin, M. M., & Czuchry, A. J. (1997). Application of root cause analysis in a service delivery operational environment: a framework for implementation. *International Journal of Service Industry Management*, 8(4), 268-289.

- [4] Li, M. H. C., Al-Refaie, A., & Yang, C. Y. (2008). DMAIC approach to improve the capability of SMT solder printing process. *IEEE Transactions on Electronics Packaging Manufacturing*, 31(2), 126-133.
- [5] Sahay, C., Ghosh, S., & Bheemarthi, P. K. (2011, January). Process improvement of brake lever production using DMAIC (+). In *ASME 2011 International Mechanical Engineering Congress and Exposition* (pp. 801-826). American Society of Mechanical Engineers.
- [6] Rudov-Clark, S. D., & Stecki, J. (2009, March). The language of FMEA: on the effective use and reuse of FMEA data. In *Sixth DSTO International Conference on Health & Usage Monitoring* (pp. 9-12).
- [7] Teoh, P. C., & Case, K. (2004). Failure modes and effects analysis through knowledge modelling. *Journal of Materials Processing Technology*, 153, 253-260.
- [8] Dimensional Metrology Standards Consortium. Part 1: Overview and fundamental principles in quality information framework (QIF) – an integrated model for manufacturing quality information, 2014.
- [9] Azar, T.A., El-Metwally, S.M. Decision tree classifiers for automated medical diagnosis. *Neutral Computing and Applications*. July 2012, 23, 2386-2403.
- [10] Riano D., Real, F., Lopez-Vallverdu, J.A., Campana, F. Ercolani, S., Mecocci, P., Annicchiarico, R., Caltagirone, C. An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients. *Journal of Biomedical Informatics*. June 2012, 45(3), 429-446.
- [11] Bouamrane, M.M., Rector, A., Hurrell, M. Using Ontologies for an Intelligent Patient Modelling, Adaption and Management System. On the Move to Meaningful Internet Systems: OTM 2008, 5332, 1458-1470.
- [12] Bouti, A., Ait, D., A State of the Art Review of FMEA-FMECA. *International Journal of Reliability, Quality and Safety Engineering*. 01(04) 1994.
- [13] Sharma, G. V. Sharma, G. V. S. S., & Rao, P. S. A DMAIC approach for process capability improvement an engine crankshaft manufacturing process. *Journal of Industrial Engineering International*, 2014. 10(2), 1-11.
- [14] Gijo, E. V. Process improvement through Six Sigma with Beta correction: a case study of manufacturing company. *The International Journal of Advanced Manufacturing Technology*, 2014. 71(1-4), 717-730.
- [15] Kaushik, P. K. A case study: Application of Six Sigma methodology in a small and medium-sized manufacturing enterprise. *The TQM Journal*, 2012. 24(1), 4-16.
- [16] Shreeranga Bhat, a. N. Application of Six Sigma methodology in a small-scale foundry industry. *International Journal of Lean Six Sigma*, 2014. 193-211.
- [17] Object Management Group (OMG). Unified Modeling Language Specification 2.5. June 2015.
- [18] ANSI/ISA-95.00.01-2010 (IEC 62264-1 Mod) Enterprise-Control System Integration - Part 1: Models and Terminology
- [19] ANSI/ISA-88.00.01-2010 Batch Control Part 1: Models and Terminology (ISA 88 document)
- [20] ISO 22400-1:2014, 2014, "Automation systems and integration—Key Performance Indicators (KPIs) for manufacturing operations management; Part 1: Overview, concepts, and terminology," International Organization for Standardization.
- [21] ISO 22400-2: 2011, "Automation systems and integration – Key performance indicators (KPIs) for manufacturing operations management – Part 2: Definitions and descriptions of KPIs," International Organization for Standardization.